SVENSKA GEOFYSISKA FÖRENINGEN

VOLUME 13, NUMBER 3 **Tellus** AUGUST 1961

A QUARTERLY JOURNAL OF GEOPHYSICS

A Mountain Wave Theory including the Effect of the Vertical Variation of Wind and Stability[1]

By BO R. DÖÖS, University of Stockholm

(Manuscript received December 24, 1960)

Abstract

The motion over an infinitely long mountain ridge has been studied. The scale of the motion is assumed to be so small that the influence of the earth's rotation can be neglected. It is assumed further that the motion is non-viscous, laminar, steady and isentropic. The resulting wave equation is solved under the assumption that the coefficient which contains the vertical variation of the wind and temperature can be approximated by the function: $\alpha_0 \exp(-\lambda z) + \alpha_1$.

The wavelengths and amplitudes are evaluated in four actual cases when lee wave clouds have been observed. The agreement between the observed and computed distances between the piles of lee wave clouds is good.

In these four cases comparison is also made with the simple method to compute the wavelength obtained by considering the oscillation of an air parcel in a stable atmosphere.

1. Introduction

There are, indeed, very few types of motion of the atmosphere which have been so much theoretically discussed as the behaviour of an air current passing over a mountain ridge. There are several possibilities to explain this fact. One reason may be for instance that this kind of motion often is clearly evident in a very beautiful display of clouds which it gives rise to and thereby invites investigation and theoretical explanation.

The classical investigations in this field are by LORD RAYLEIGH (1883) and LORD KELVIN (1886) who studied standing waves on a stream of water. LYRA (1943) and QUENEY (1947, 1948) gave solutions for a stratified uniform air current but the first attempt to introduce the effects of both the variation of wind and stability with height was by SCORER (1949). Since then the most important contributions seem to have been the solutions of the transient problem which HÖILAND (1951) has solved for surface waves and which PALM (1953) and WURTELE (1953, 1955) succeeded in solving for the same model used by Lyra and Queney. In these solutions it is shown that the initial value solution approaches the stationary solution. The main advantage of treating it as an initial value problem is that the unique solution is obtained without introducing any artificial methods. On the other hand, such artificiality is necessary if the motion is assumed to be stationary and free waves exist.

Most of the theoretical work in this field has been concerned with highly simplified assumptions regarding the vertical variation of the wind and the stability. One exception is to be mentioned here namely SAWYER (1960) who solved the lee-wave equation numerically with the aid of an electronic computer.

[1] The research reported here has been sponsered in part by the Geophysics Research Directorate, AFCRL, AFRD, of the Air Research and Development Command, USAF, through its European office.

The purpose of the present study is to compare observations of lee-waves with results obtained theoretically with the aid of a method, discussed previously by the author (1958). This method permits the stability and the temperature to vary with height in a realistic way. In this connection it should be mentioned that a similar method of solution to the lee-wave equation has been used independently by PALM and FOLDVIK (1959, 1960) in an investigation of lee-waves in the stratosphere.

2. The lee-wave equation

The derivation of the lee-wave equation will not be given in more detail than needed to make clear the various assumptions and approximations which are necessary in order to obtain it in a sufficiently simple form.

Since the scale of the motion in which we are interested is relatively small (the wave length is of the order of magnitude 10 km) we can quite safely disregard the effect of the earth's rotation. We further assume that the amplitude of the waves is relatively small compared to the wavelength so that we can neglect the non-linear terms in the basic equations. This assumption is justified in the cases with which we will be concerned.

The motion will be studied in a coordinate system with the z-axis upwards and the x-axis directed along the undisturbed current U, which is assumed to be a function of height only. The airstream blows perpendicularly to the mountain ridge which is considered to have an infinite extension along the y-axis so that the motion is assumed to be two-dimensional in the xz-plane. We assume further that the motion is nonviscous, laminar and isentropic. The temperature and the wind speed of the undisturbed current are allowed to vary with height.

For this type of motion it would be too complicated to deal with an initial value problem. However, it is reasonable to expect that also here the transient solution should approach the stationary solution.

The basic equations which govern the motion are then: two equations of motion, the equation of continuity, the equation of state and the equation of piezotropy. The last equation expresses the dry adiabatic changes of state.

After linearization and elimination we arrive at the following equation for the vertical velocity w':

$$\left(1-\frac{U^2}{c_0^2}\right)\frac{\partial^2 w'}{\partial x^2}+\frac{\partial^2 w'}{\partial z^2}-$$

$$-\left(\frac{\varkappa g}{c_0^2}-\frac{\varkappa R\gamma+2U\dfrac{dU}{dz}}{c_0^2-U^2}\right)\frac{\partial w'}{\partial z}+$$

$$+\left[\frac{(\varkappa-1)g^2}{c_0^2-U^2}-\frac{\varkappa Rg\gamma}{U^2(c_0^2-U^2)}-\frac{1}{U}\frac{d^2U}{dz^2}+\right.$$

$$+\left(\frac{\varkappa g}{c_0^2}-\frac{\varkappa R\gamma+2g}{c_0^2-U^2}\right)\frac{1}{U}\frac{dU}{dz}-$$

$$\left.-\frac{2}{c_0^2-U^2}\left(\frac{dU}{dz}\right)^2\right]w'=0 \qquad (1)$$

where

$$c_0^2=\varkappa RT_0$$

The subscript zero refers to the undisturbed current and the primes to the disturbed flow produced by the mountain ridge. γ is the lapse rate of the undisturbed current, c_0 the velocity of sound, $\varkappa$ the ratio of the specific heat at constant pressure and constant volume, g the acceleration of gravity and R the gas constant for unit mass.

In equation (1) we can with very good approximation neglect U^2 in comparison with c_0^2. The coefficient for $\partial w'/\partial z$ can then be written:

$$\frac{g-R\gamma}{RT_0}-\frac{2U}{\varkappa RT_0}\frac{dU}{dz}$$

In none of the observed cases which will be compared with the theory, does the product UdU/dz exceed 100 $m^2\ sec^{-2}\ km^{-1}$, and the lapse rate is always less than 10° C km^{-1}. For these extreme values the ratio between the first and second term is about 50. We can therefore with good approximation disregard the second term. We introduce now the new dependent variable w_1, defined by

$$w'=w_1\exp\left\{\frac{g-R\gamma}{2RT_0}z\right\} \qquad (2)$$

Since the quantity

$$\frac{g-R\gamma}{2RT_0}=\frac{34.2-\gamma}{2T_0}\,\mathrm{km}^{-1}$$

varies relatively little with height we can disregard its variation when we differentiate (2) in the course of changing variables. Equation (1) becomes then

$$\frac{\partial^2 w_1}{\partial x^2} + \frac{\partial^2 w_1}{\partial z^2} + \alpha(z) w_1 = 0 \qquad (3)$$

where

$$\alpha(z) = \frac{g(\gamma_d - \gamma)}{U^2 T_0} - \frac{1}{U}\frac{d^2U}{dz^2} + \\ + \left(\frac{\gamma_d - \gamma}{T_0} - \frac{g}{\varkappa R T_0}\right)\frac{1}{U}\frac{dU}{dz} - \\ - \frac{2}{\varkappa R T_0}\left(\frac{dU}{dz}\right)^2 - \left(\frac{g - R\gamma}{2RT_0}\right)^2 \qquad (4)$$

and γ_d is the dry adiabatic lapse rate.

As a lower boundary condition we require that the flow is tangential to the surface. If $z = \zeta_s(x)$ is the equation for the mountain profile we have thus

$$w'(x, 0) = w_1(x, 0) = U(\zeta_s)\frac{\partial \zeta_s}{\partial x} \qquad (5)$$

We further require that the energy of the waves remains finite at great heights. In order to fulfill this requirement we use as the upper boundary condition

$$w_1(x, z) = 0 \quad \text{when } z \to \infty$$

To solve the equation (3) we follow the usual procedure of assuming that the ground profile is sinusoidal, and thereafter we generalize the solution by Fourier integrals for an arbitrary mountain.

We then assume that we can write

$$w_1 = We^{ikx} \qquad (6)$$

where k is the wave number.

Regarding the coefficient $\alpha(z)$ (4) which contains the vertical variation of the stability and the wind we will see later that this can be represented with good approximation, at least in the troposphere, by the function:

$$\alpha^*(z) = \alpha_0 e^{-\lambda z} + \alpha_1 \qquad (7)$$

where α_0, α_1, and λ are parameters we have to determine from observations.

With (6) and (7) equation (3) can be written:

$$\frac{d^2W}{dz^2} + (\alpha_0 e^{-\lambda z} + \alpha_1 - k^2)W = 0 \qquad (8)$$

This equation can be transformed to a Bessel equation by the change of the independent variable

$$\eta = \beta e^{-\frac{\lambda z}{2}} \qquad (9)$$

where

$$\beta = 2\frac{\sqrt{\alpha_0}}{\lambda} \qquad (10)$$

(8) then becomes

$$\eta^2 \frac{d^2W}{d\eta^2} + \eta\frac{dW}{d\eta} + \left[\eta^2 - \frac{4}{\lambda^2}(k^2 - \alpha_1)\right]W = 0 \qquad (11)$$

The solution to this equation after introducing (9) can be written:

$$W = AJ_\nu\left(\beta e^{-\frac{\lambda z}{2}}\right) + BY_\nu\left(\beta e^{-\frac{\lambda z}{2}}\right) \qquad (12)$$

where $\nu = \frac{2}{\lambda}\sqrt{k^2 - \alpha_1}$ for $k \geqslant k_0 = \sqrt{\alpha_1}$

$$W = CJ_{i\mu}\left(\beta e^{-\frac{\lambda z}{2}}\right) + DJ_{-i\mu}\left(\beta e^{-\frac{\lambda z}{2}}\right) \qquad (13)$$

where $\mu = \frac{2}{\lambda}\sqrt{\alpha_1 - k^2}$ for $k < k_0 = \sqrt{\alpha_1}$

Since Bessel functions of the second kind with real order approach minus infinity when the argument tends to zero ($z \to \infty$), we must put $B = 0$ in order to fulfill the upper boundary condition.

For wave numbers less than $\sqrt{\alpha_1}$ the solution (13) is given in terms of Bessel functions with imaginary order. The behaviour of these is very little discussed in the literature. It is therefore preferable to use the solution BOOLE (1844) obtained for the Bessel equation with the order $i\mu$

$$W = \cos(\mu \log \eta)(a_0 + a_2\eta^2 + a_4\eta^4 + \ldots) \\ + \sin(\mu \log \eta)(b_0 + b_2\eta^2 + b_4\eta^4 + \ldots)$$

where in general

$$\begin{cases} a_{2n} = -\dfrac{n a_{2n-2} - \mu b_{2n-2}}{4n(n^2 + \mu^2)} \\ b_{2n} = -\dfrac{n b_{2n-2} + \mu a_{2n-2}}{4n(n^2 + \mu^2)} \end{cases} \qquad (14)$$

Making use of (9) we have thus the solution for $k < k_0 = \sqrt{\alpha_1}$

$$W = \cos\left(\mu \log \beta - \frac{\mu\lambda z}{2}\right) \sum_{n=0}^{\infty} a_{2n}\beta^{2n} e^{-n\lambda z}$$

$$+ \sin\left(\mu \log \beta - \frac{\mu\lambda z}{2}\right) \sum_{n=0}^{\infty} b_{2n}\beta^{2n} e^{-n\lambda z} \quad (15)$$

where a_{2n} and b_{2n} are given by (14).

When z goes to infinity this solution will be trigonometric. The upper boundary condition cannot be satisfied unless both a_0 and b_0 are put equal to zero. The solution is therefore identically zero for wave numbers less than $k_0 = \sqrt{\alpha_1}$. This indicates that the upper boundary condition is quite severe. If we instead use the less restrictive upper boundary condition that w_1 remains finite when z goes to infinity, the constants a_0 and b_0 could be determined, so that both this condition and the lower boundary condition have been satisfied. However, the contribution to the Fourier integral for values of k in this interval ($0 \leq k < k_0$) would have been very small. It can be shown that there are no free waves in this interval. Regardless of which one of these two versions of the upper boundary condition we chose we can thus safely disregard the contribution to the Fourier integral in the interval $0 \leq k < k_0$.

From (12) together with (2) and (6) we have then the following solution for the vertical velocity:

$$w'(x,z) = Ae^{ikx} \exp\left\{\frac{g - R\gamma}{2RT_0} z\right\} J_\nu\left(\beta e^{-\frac{\lambda z}{2}}\right) \quad (16)$$

for $$k \geqslant k_0 = \sqrt{\alpha_1}$$

This particular solution of the wave equation will now be used to obtain the solution for the single mountain ridge. The profile we will adopt is the one QUENEY (1948) used, namely

$$\zeta_s(x) = \frac{b}{1 + \left(\frac{x}{a}\right)^2} = ab \int_0^\infty e^{-ak} \cos kx dk \quad (17)$$

where a denotes the half width of the mountain and b the height at the crest. For this mountain the linearized lower boundary condition is

$$w'(x,0) = U(0) \frac{\partial \zeta_s(x)}{\partial x}$$

$$= U(0)\, ab \frac{\partial}{\partial x} \int_0^\infty e^{-ak} \cos kx dk \quad (18)$$

The solution satisfying this condition is:

$$w'(x,z) = U(0) ab \exp\left\{\frac{g - R\gamma}{2RT_0} z\right\}$$

$$\frac{\partial}{\partial x} \int_{k_0}^\infty e^{-ak} \cos kx \frac{J_\nu\left(\beta e^{-\frac{\lambda z}{2}}\right)}{J_\nu(\beta)} dk \quad (19)$$

After integration we obtain the displacements of the streamlines,

$$\zeta(x,z) = \frac{U(0)}{U(z)} ab \exp\left\{\frac{g - R\gamma}{2RT_0} z\right\}$$

$$\int_{k_0}^\infty e^{-ak} \cos kx \frac{J_\nu\left(\beta e^{-\frac{\lambda z}{2}}\right)}{J_\nu(\beta)} dk \quad (20)$$

which also may be written

$$\zeta(x,z) = F(z)\, Re \int_{k_0}^\infty e^{-ak} e^{ikx} \frac{J_\nu\left(\beta e^{-\frac{\lambda z}{2}}\right)}{J_\nu(\beta)} dk \quad (21)$$

where

$$F(z) = \frac{U(0)}{U(z)} ab \exp\left\{\frac{g - R\gamma}{2RT_0} z\right\} \quad (22)$$

In order to evaluate the integral (21) we apply Cauchy's residue value theorem. We therefore consider the integral

$$I = F(z) \int_C e^{-ak} e^{ikx} \frac{J_\nu\left(\beta e^{-\frac{\lambda z}{2}}\right)}{J_\nu(\beta)} dk \quad (23)$$

where C is the path of integration in the complex k-plane (fig. 1). Throughout this integration the order of the Bessel function is denoted by ν, which thus here can assume complex values.

According to COULOMB (1936) the zeros of $J_\nu(\beta)$, regarded as a function of the order ν, are real and simple for positive values of the argument β. The integrand has therefore poles only

for real values of $k \geqslant k_0 = \sqrt{\alpha_1}$. We denote these by $k_1, k_2 \ldots k_N$.

The path of integration has thus indentations along the real axis. The radii of these are denoted $\varepsilon_1, \varepsilon_2, \ldots \varepsilon_N$. If x is positive the contour is in the first quadrant, and if x is negative it is in the fourth. The integrand is regular within and on C; I is thus zero. Evaluation of the various parts of the integral gives after making $\varepsilon_1, \varepsilon_2, \ldots \varepsilon_N \to 0$.

$$0 = F(z) P \int_{k_0}^{R} e^{-ak + ikx} \frac{J_\nu\left(\beta e^{-\frac{\lambda z}{2}}\right)}{J_\nu(\beta)}\, dk$$

$$\mp F(z)\, \pi i \sum_{n=1}^{N} Res(k_n) + F(z)\left\{I_{A1,2} + I_{B1,2}\right\} \quad (24)$$

where P denotes the Cauchy principal value. The second term is the contribution along the indentations at the poles. The upper sign applies when x is positive and the lower when x is negative. $I_{A1,2}$ represents the integrals along the circular axis. It can be shown that these integrals vanish when the radius R goes to infinity. The last integrals $I_{B1,2}$ represent the integrals along the lines parallel to the imaginary axis. By taking real parts of (24) and after evaluating the residuals we obtain with reference to (21)

$$\zeta(x, z) = \mp \pi F(z)$$

$$\sum_{n=1}^{N} \left[e^{-ak} \sin kx \frac{J_\nu\left(\beta e^{-\frac{\lambda z}{2}}\right)}{\frac{d}{dk} J_\nu(\beta)} \right]_{k=k_n} - ReF(z) I_{B1,2} \quad (25)$$

The last term in this solution is only important in the close vicinity of the mountain; its value decreases with increasing $|x|$ at a rate of $e^{-2\pi}$ per wave length. Since we are primarily interested in the waves set up by the mountain we will not here make any attempt to evaluate this term. Regarding the first N terms it can be seen that they give waves with different wave lengths both upstream and downstream. The solution is not unique for as mentioned in the introduction this difficulty arises when the motion is assumed to be stationary and free waves exist. Different methods have been constructed to avoid this. RAYLEIGH (1883) showed that the problem could be uniquely determined by introducing small frictional forces proportional to the velocity. In the final solution these forces were put equal to zero. KELVIN (1886) devised

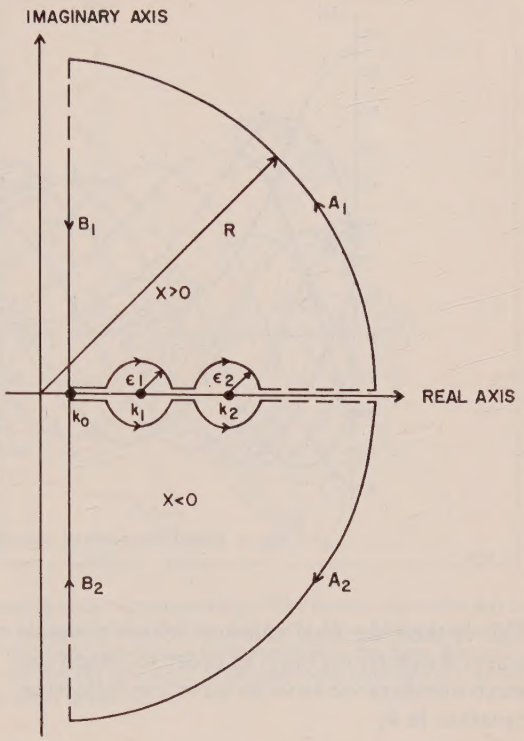


Fig. 1. Paths for evaluation of the integral (23)

another method to obtain the unique solution, the so-called monotony condition, which we will make use of. A third method is Sommerfeld's "Ausstrahlung" condition. This method was rejected by LYRA (1943). ZIEREP (1957), however, showed that it can be used. ELIASSEN and PALM (1954) have demonstrated that the stationary solution is made unique by requiring the mountain to be the only energy source.

Lord Kelvin's monotony condition consists, in short, of adding free waves with such amplitudes and wave lengths that the disturbances vanish for large negative x. This gives in our case

$$\zeta(x, z) = -2\pi F(z) \sum_{n=1}^{N} \left[e^{-ak} \sin kx \frac{J_\nu\left(\beta e^{-\frac{\lambda z}{2}}\right)}{\frac{d}{dk} J_\nu(\beta)} \right]_{k=k_n} - ReF(z) I_{B1}: \quad x > 0$$

$$\zeta(x, z) = -ReF(z) I_{B2}: \quad x < 0 \quad (26)$$

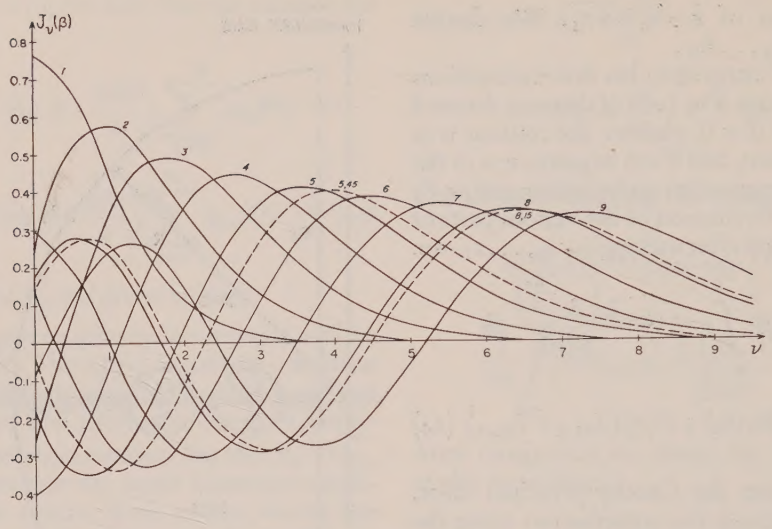


Fig. 2. Bessel functions of constant argument and variable order.

This is then the final solution which contains waves downstream only. In order to obtain the wave numbers we have to solve the following equation in k.

$$J_\nu(\beta) = 0 \tag{27}$$

The roots $k_1, k_2 \ldots k_N$ (poles of the integrand of (23)) are easily found graphically. In fig. 2 is the Bessel function shown as a function of ν (from JAHNKE-EMDE, 1945). From the positive roots $\nu_1, \nu_2 \ldots \nu_N$ the wave numbers are obtained from the relation between ν and k.

$$k_n = \sqrt{\alpha_1 + \left(\frac{\lambda \nu_n}{2}\right)^2} \tag{28}$$

The wave lengths are thus

$$L_n = \frac{2\pi}{\sqrt{\alpha_1 + \left(\frac{\lambda \nu_n}{2}\right)^2}} \tag{29}$$

The amplitudes of these waves are given by

$$A_n = \lambda \pi a b \frac{U(0)}{U(z)} \exp\left\{\frac{g - R\gamma}{2RT_0} z - \right.$$

$$\left. - a k_n \right\} \frac{\sqrt{k_n^2 - \alpha_1}}{k_n} \left[\frac{J_\nu\left(\beta e^{-\frac{\lambda z}{2}}\right)}{\frac{d}{d\nu} J_\nu(\beta)} \right]_{k=k_n} \tag{30}$$

Both the numerator and the denominator in the last factor can then be graphically evaluated with the aid of fig. 2.

In the next section we will make use of observations of lee wave clouds in order to obtain an estimate of the accuracy with which the wavelength can be computed with the aid of this method. The behaviour of the theoretically obtained streamlines in the vicinity of the mountain could unfortunately not be compared with the available observations. Therefore the term in the solution which is of importance in that region will not be evaluated.

3. Comparisons between observed and theoretical results

Field studies of lee wave clouds have been carried out in various parts of the world, e.g. KÜTTNER (1939), FÖRCHTGOTT (1949), DE VER COLSON (1952). In the Ovik mountains in Sweden very detailed and carefully undertaken observations have been performed by LARSSON (1954, 1958), from which the following information is obtained.

The form and position of these mountains (cf. fig. 3) have been found to be exceedingly favorable for development of lee waves. As can be expected from the orientation of the mountains lee waves will most likely occur for SW and NW winds. The profile of the mountain

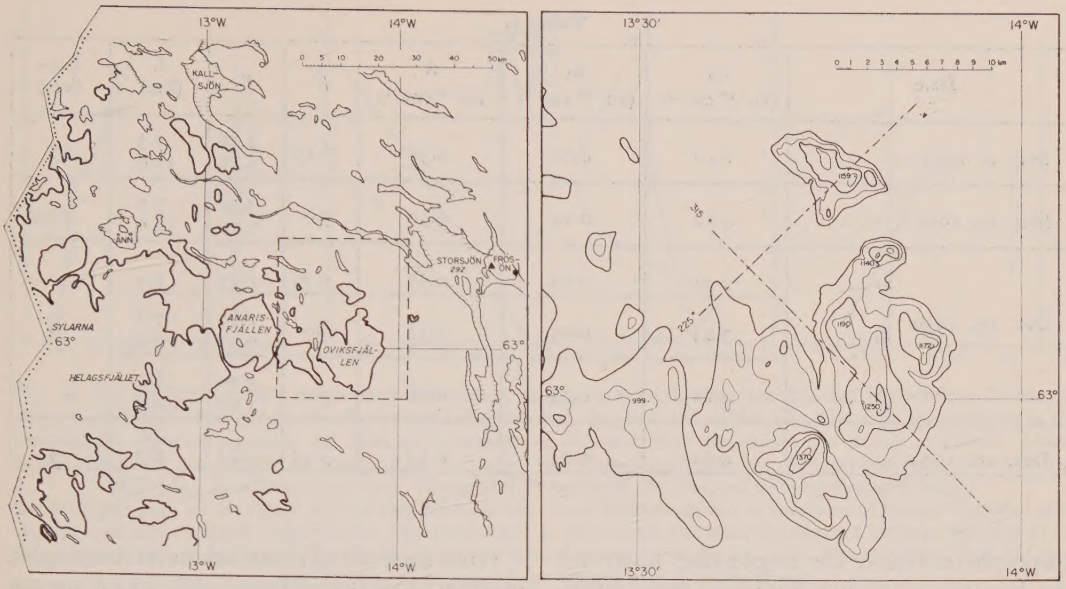


Fig. 3. Map of the eastern side of the central part of the Scandinavian mountain range. The broken line to the left is the Norwegian border. The triangle indicates the observation place at Frösön and the dot the position of the radiosonde station (02-062). The Ovik mountains (inside the dashed rectangle) are shown more detailed to the right. Contour-lines are drawn for each 100 m from 800 m. In this figure are also the two main wind directions indicated for which lee waves most often occur.

along the two directions indicated in fig. 3 agrees fairly well with the theoretically assumed profile given by equation (17). The assumption made before, namely that the mountain range has an infinite extension perpendicular to the main current, is of course not true but this deficiency is not likely to be too severe. The height b over the surrounding country, which is fairly level, is about 700 metres for both directions. The other mountain parameter a was chosen to be 3 km for NW-wind situations and 2.2 km for SW-wind situations.

Wind and temperature distributions were obtained from the radiosonde station at Frösön (02-062) which is situated about 40 kilometres to the NE of the Ovik mountains.

For the comparison of the theoretically computed values of the wave lengths with the observed wavelengths a number of different cases have been selected. The principal basis for the selection of the suitable observational material, from the several hundred available cases, has been that the lee wave clouds were well defined. At least two piles of clouds in the lee of the mountain were required in order to obtain a sufficiently accurate estimate of the wavelength. Furthermore, the cases chosen represented different types of vertical temperature distribution. Of course it is also important that the measurements of the clouds were undertaken at a time which was relatively close to the time of the sounding and we will discuss here four cases fulfilling these criteria.

By using the observed values for the temperature and the wind the coefficient $\alpha(z)$ (4) was evaluated as a function of height. The parameters α_0, α_1 and λ were then determined so that the function $\alpha^*(z)$ (7) agreed as well as possible with $\alpha(z)$, (cf. Table I).

The wind and temperature distributions with height for these four cases are shown in the figures 4—7. The observed wind profile is given by the dashed curves while the observed temperatures are given by dots and the dew point temperatures by circles. The solid lines represent the idealized wind and temperature distribution which permits us to replace the coefficient $\alpha(z)$ by $\alpha^*(z)$. As can be seen from the figures the observed and idealized variations of wind and temperature agree fairly well in the

Table 1.

Date	α_0 (10^{-10} cm^{-2})	α_1 (10^{-10} cm^{-2})	λ (10^{-5} cm^{-1})	β	ν	L (km)	L_{obs} (km)
May 9, 1955	2.40	0.02	0.38	8.15	4.48 1.81	7.3 16.9	15
May 19, 1959	4.12	0.10	0.58	7.0	3.50 1.00	5.9 14.6	13.5
Dec. 15, 1959 a)	2.42	0.05	0.58	5.36	2.17	9.4	8
Dec. 15, 1959 b)	3.03	0.05	0.31	6.0	2.67 0,31	7.8 26.0	8
Dec. 15, 1959 c)	4.12	0.03	0.58	7.0	3.50 1.00	6.0 18.6	8
Dec. 16, 1959	2.25	0.04	0.55	5.45	2.24	9.7	9

troposphere. Above the tropopause, however, the function $\alpha^*(z)$ does not give a good representation. On the other hand it has been shown that conditions in these regions have little influence on the motion in the middle and lower troposphere. (PALM and FOLDVIK, l.c.)

May 9, 1955.

The lee wave clouds persisted during a considerable long period of time, 0730—0820, 1245—1900, GMT. The measurements of the clouds were undertaken at approximately 1400 GMT when they assumed their most regular shape. The estimation of the altitudes of the clouds, their sizes and their positions were based on visual observations from Frösön with the aid of theodolites and on time lapse films. In this case observations were also made at the same time from the air by a pilot from the Swedish Air Force wing at Frösön.

Two piles of lee wave clouds were observed. The first one, 13 km from the crest of the mountain extended in the vertical from about 2,000 to 6,000 metres. In the horizontal it was very wide, about 60 km, and had an arch-shaped form around the Ovik mountains. Its dimension in the direction of the main current was 6 km. The second pile, however, was considerably smaller. It was oriented 28 km to the NE of the mountain and had a vertical extension of about 1,000 metres with the cloud base at about 2,000 metres. Also in the horizontal it was much smaller than the first pile, only 15 km wide.

From these cloud observations it is reasonable to expect that the accuracy with which we can estimate the wave length is about 1 km.

Unfortunately on this day the radiosonde balloon reached to only 6 km. Above this level the wind was estimated with the aid of upper air maps (the dotted line). The temperature could be obtained with the aid of previous and following soundings since there was very little variation of temperature with time during this period. The static stability did not show any marked variation with height in the troposphere. The average lapse rate was 6.6° C per km.

It is seen from fig. 4a that the amplitudes for both solutions are of importance in the lower layers. In order to make it possible to compare the theoretical results with the observations of the clouds it is therefore necessary to take into account both solutions. The streamlines at the levels 2, 3 and 4 km have been constructed and are shown in fig. 4b. At the height 2 km, which is approximately the level of the base of the lee wave clouds, the peaks of the streamline almost exactly coincides with the positions of the lee wave clouds. As a matter of fact the agreement is much better than one could expect taking into account the accuracy with which it is possible to measure the orientation of the clouds. There is of course also an uncertainty in the theoretically computed results since these also are based on observations.

Higher up the agreement is less good. The relatively great vertical extent of the first pile

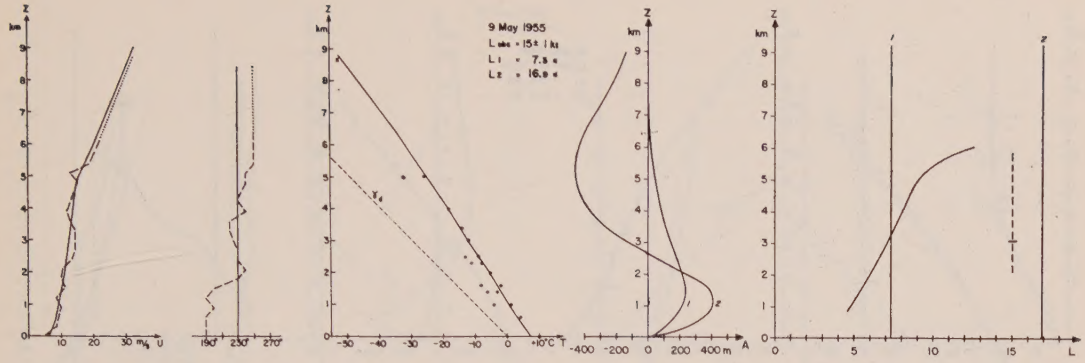


Fig. 4a. To the left is shown the vertical distribution of the wind speed, wind direction and temperature for May 9, 1955, 1500 GMT. The dashed line in the wind diagram and the dots and circles (the dewpoint temperature) in the temperature diagram represent the observed conditions as obtained from the radiosonde station at Frösön. Above 6 km the wind was estimated from upper air maps (the dotted line) and the temperature from previous and following soundings (the cross). The solid lines represent the idealized distributions with height.
In the right part of the figure are the amplitudes of the wave solutions given as functions of height. Together with the theoretically obtained wavelengths (the solid lines) is also shown the observed distance between the piles of lee wave clouds (the dashed line). This line also gives the vertical extension of the first pile. The horisontal dash indicates the height to the top of the second pile of wave clouds.
The slanting solid line is the vertical variation of the wavelength computed with the aid of the formula (31).

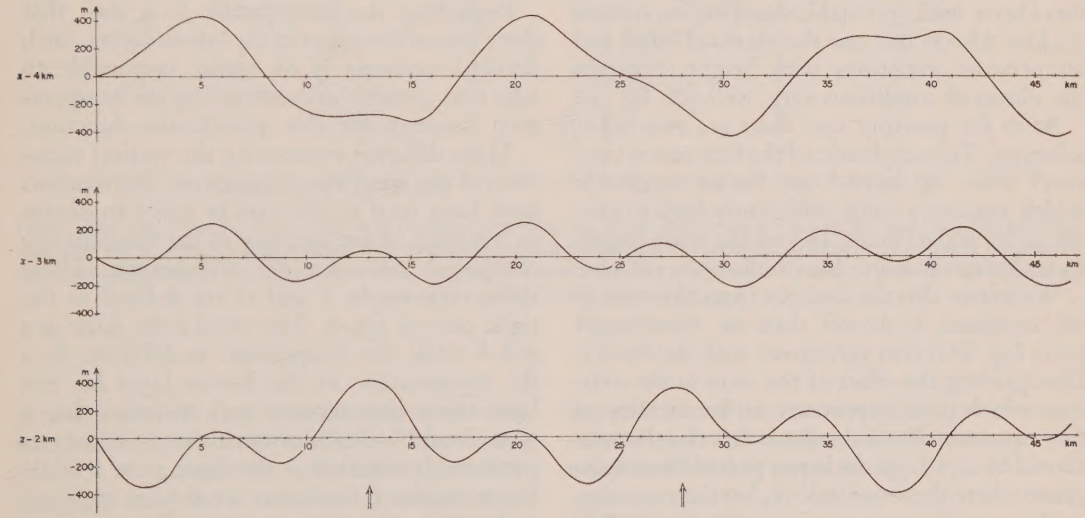


Fig. 4b. Streamlines at the levels $z = 3$, 4 and 5 km in the case May 9, 1955. The two arrows at the bottom of the figure indicates the position of the two piles of clouds observed in this case.

cannot be explained. In this connection it is to be remembered that the term in the solution (27) which is of importance close to the mountain is not taken into account here. We may also think of the possibility that the heat released in the cloud could create sufficiently strong vertical motion so that the streamlines higher up will be changed. This "secondary effect" deserves investigation.

May 19, 1959.

The record of observations on this day says that lee wave clouds occurred between 1300 and 1700 GMT. Three piles of lee wave clouds were observed SE of the Ovik mountains. The base of the piles were at about 1,700 metres. The first pile had a great vertical extent, now and then up to 5 km. The center of this pile was oriented about 8.5 km from the crest of the

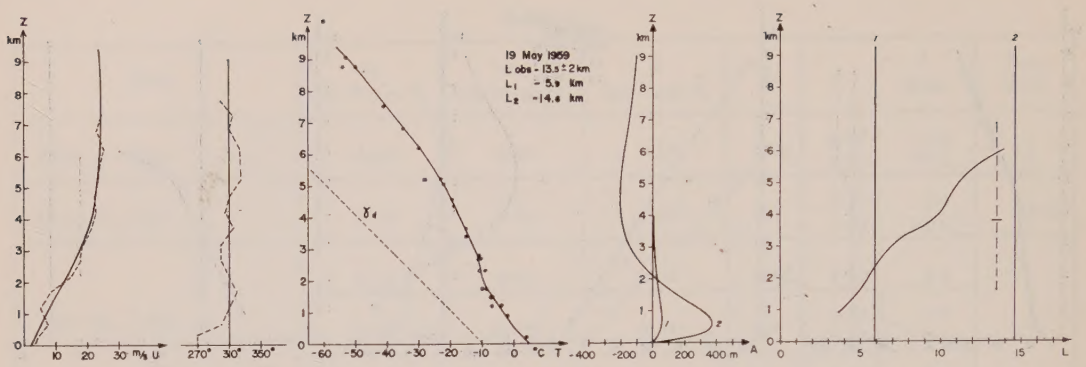


Fig. 5. The case May 19, 1959, 1200 GMT. (For further explanation see fig. 4a.)

mountain. The two other piles were considerably less thick, 1—3 km. Below the lee wave clouds there was strato-cumulus and above there were cirrus bands with ripples.

The wind observations reached to about 7.5 km. The temperature curve shows that there was a layer with very stable stratification around 2.5 km. Also in this case the idealized wind and temperature variations with height represent the observed conditions very well (cf. fig. 5).

As in the previous case there are two wave solutions. The amplitude of the first one is very small while the second one has an amplitude which assumes a value sufficiently high to give rise to lee wave clouds. Indeed the wave length L_2 is also much more close to the observed one.

We notice that the distance from the crest of the mountain is shorter than the wavelength (0.63 L_2). This is in agreement with the theory. Disregarding the effect of the term in the solution which is of importance in the vicinity of the mountain it is easy to see that this distance should be 0.75 L_2 in the lower part of the atmosphere where the amplitude A_2 has the same sign as A_1.

As in the previous case the tops of the piles, particularly the first one, are higher than can be expected from the vertical variation of the amplitude.

December 15, 1959.

Unfortunately there was no pibal this day at Frösön. However, the upper air flow did not show any strong horizontal variations over eastern Scandinavia. The wind at the Ovik mountains could therefore reasonably well be estimated from the wind measurements at the surrounding stations. Therefore in fig. 6 the wind variation at 1200 GMT are shown for the three stations Bodö (01-052), Orlandet (01-241) and Gardermoen (01-384). As before the solid line represents the idealized wind distribution with height.

Regarding the temperature it is seen that there was an inversion in the lowest layers. Such detailed variation is of course impossible to take into account in constructing the temperature variation for the approximate function.

Three different versions for the vertical variations of the wind and temperature distributions have been used in this case in order to obtain an estimate of the accuracy of the theoretically computed value for the wavelength. These three versions (*a*, *b* and *c*) are defined in the right part of fig. 6. The wind is the same in *a* and *b* while the temperature is different. In *a* the temperature in the lowest layer has not been taken into account and therefore has a more unstable stratification there. In *c* the temperature distribution is the same as in *b* while the strength of the mean wind U is approximately 2 m/s less.

The most realistic choice of these three versions is *a* regarding both the wind and the temperature. The first pile in the lee of the mountain was situated at a distance of 10 km from the crest and the base was rather low, about 1,200 metres. The size of the wave clouds in the direction of the main current increased with height, thereby giving the pile a comet-like shape. The second pile was observed 8 km downstream from the first one with its base at 2,000 m. Higher up the shape of the wave clouds in this pile were not very well defined. Still fur-

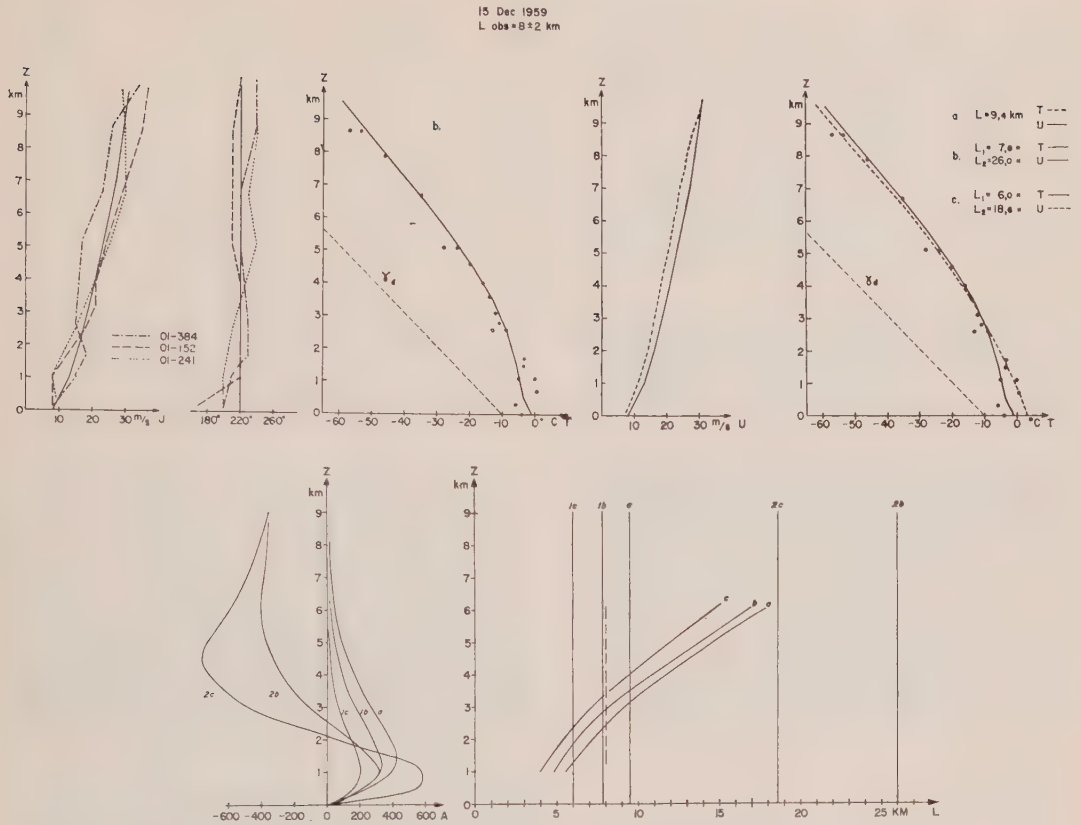


Fig. 6. The case Dec. 15, 1959, 1200 GMT. Since there were no wind observations this day the wind was interpolated with the aid of surrounding stations. The wavelengths and the amplitudes are evaluated for three different versions of the wind and temperature distribution with height; two with different temperature (a and b) and two with different winds. (For further explanation see fig. 4a.)

ther downstream there were irregularly oriented lenticularis clouds at various levels. The version *a* which was considered to be most representative gives only one wave solution with a wavelength of 9.4 km. With the aid of the other two versions it is possible to obtain an estimate of the accuracy of the solution. It is seen that the wavelength for the first solution (L_1) is much less dependent on changes in the wind and the stability compared with the second solution (L_2).

The difference in the wavelength L_1 for the two versions *b* and *c* with different wind distribution with height is 1.8 km while the corresponding difference for the two versions (*a* and *b*) with different stability distribution is 1.6 km. These differences are to be considered as the upper limits of the inaccuracy in L_1 due to uncertainties in the idealized distributions for the wind and stability respectively.

December 16, 1959.

On this day the lenticularis clouds which were observed had been present continuously since the previous day. The wind had not changed appreciably over night, the direction remaining practically the same while the speed decreased somewhat. The stratification was slightly less stable on this day (approximately moist adiabatic) and the height of approximately 1 km there was an isothermal layer which has not been taken into account. The moisture was relatively high this day, particularly in the layer 3 to 5 km (cf. fig. 7).

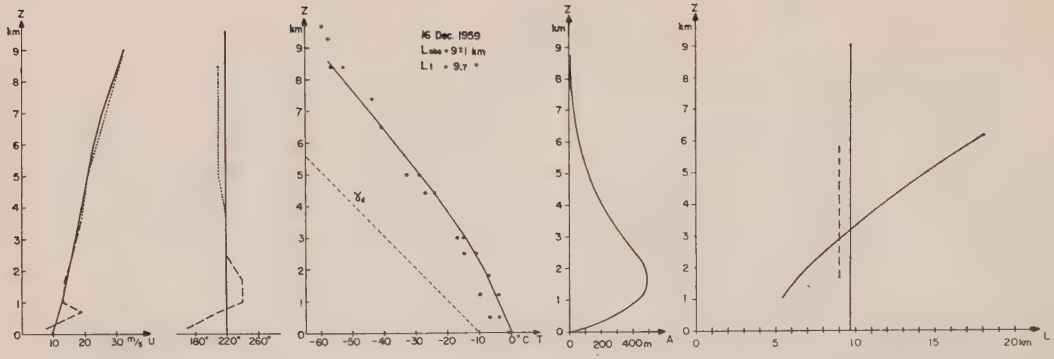


Fig. 7. The case Dec. 16, 1959, 1200 GMT. (For further explanation see fig.4a.)

Three piles of lee wave clouds were observed around the time of the sounding at 1200 GMT. The two first piles extended approximately from 2 to 6 km and below the two first piles were found rotor clouds. The third one had a base at about 4 km, whereas higher up it was less clearly defined. The distances to the three piles from the crest of the mountain were respectively 8.17 and 26 km. In this case the theory gave only the one wavelength 9.7 km.

4. Discussion

The theoretical method we have applied here seems to give good agreement with observed cloud pattern for waves in the lee of a mountain of such a dimension that the effect of the earth's rotation can be neglected. The agreement between the theoretically computed and the observed wavelength is well within the limits of the uncertainty in the observations of the distances between the piles of lee wave clouds. The influence of the variation in the computed wavleength due to inaccuracies in the idealized representations of the observed wind and temperature has been investigated in one case only (December 15, 1959). This was, however, the case where these idealized distributions probably were least representative. With this in mind it is reasonable to expect that the uncertainty in the computed wavelength could be estimated to be about 1 km in the other three cases.

Regarding the amplitude it is more difficult to make comparisons with the observations. However, there is no doubt that the accuracy of the computed amplitude shows obvious deviations from what one can obtain from the observations of the clouds. Since these deviations are most marked in the upper part of the troposphere it is natural to expect that they are a consequence of the inaccuracy in the representation of the observed wind and temperature above the tropopause. Another possible explanation has been suggested before namely that the heat released due to condensation in the wave clouds could increase the vertical motion and thereby increase the vertical extent of the piles.

As we have seen the solution for the streamlines contains a number of waves superimposed on each other. This number is given by the number of roots to the equation (27) and is determined by the value of the parameter β (cf. fig. 2) which in turn depends on the vertical variation of $\alpha(z)$, i.e. on the wind and stability. From Jahnke Emde (1948) we obtain

$$0 \leq \beta < 2.405 \text{ no waves}$$
$$2.405 \leq \beta < 5.520 \text{ one wave}$$
$$5.520 \leq \beta < 8.654 \text{ two waves}$$
$$8.654 \leq \beta < 11.792 \text{ three waves}$$

. . . .
. . . .
. . . .

In the atmosphere the wind and temperature conditions are usually such that one or two waves exist.

Since for constant α_0 the parameter β decreases with increasing λ we can conclude than if $\alpha^*(z)$ decreases sufficiently rapidly with height β becomes less than 2.405 and no waves exist. In fig. 8 is shown a hypothetical case where the

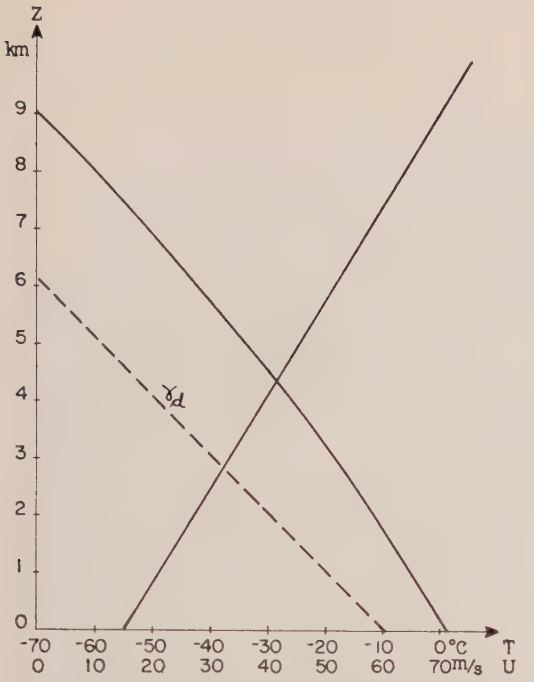


Fig. 8. Wind and temperature distributions with height in a hypothetical case for which no waves occur. $\alpha_0 = 0.52$, $\alpha_1 = 0$, $\lambda = 0.6$, $\beta = 2.404$.

wind and temperature have such a distribution that no waves occur. The values of the parameters were for this case:

$\alpha_0 = 0.52$, $\alpha_1 = 0$ and $\lambda = 0.6$, then $\beta = 2.404$. The stability varies relatively little with height and in the lower layers it is approximately a moist adiabatic lapse rate. With increasing height the stability decreases and becomes almost a dry adiabatic lapse rate at about 8 km. This type of temperature curve is quite realistic but the wind profile is more extreme; the increase with height is relatively strong. From the way the wind and stability occur in the coefficient $\alpha(z)$, (4), we can draw the conclusions that a case where the windspeed is less or if the stability is stronger, α_0 is greater and consequently also β. Then waves can occur. Similarly, if the wind or the lapse rate increases less rapidly with height, λ becomes smaller and β increases. In the limiting case ($\beta = 2.405$) the wavelength (29) is infinitely long (provided $\alpha_1 = 0$) and it decreases with increasing β.

A drawback with the method used here is that the vertical variation of the wind and stability must be approximated with an analytical function. It is quite clear from the preceding discussion that this is possible in many cases. Of course it has been necessary to smooth the curves representing the observed conditions but this is usually not serious. On the contrary it is to be recommended since it would be false accuracy to take all the details in the curves into account. However, there are cases for which it has been impossible to represent the observed conditions in this way.

In the observational study of LARSSON (1954) the stability conditions when lee wave clouds occurred have been investigated. The mean temperature curve when lee waves occurred was constructed from 19 soundings during March 1953 (cf. fig. 9). This is compared with the mean temperature for the same month when lee wave clouds did not occur (43). A consider-

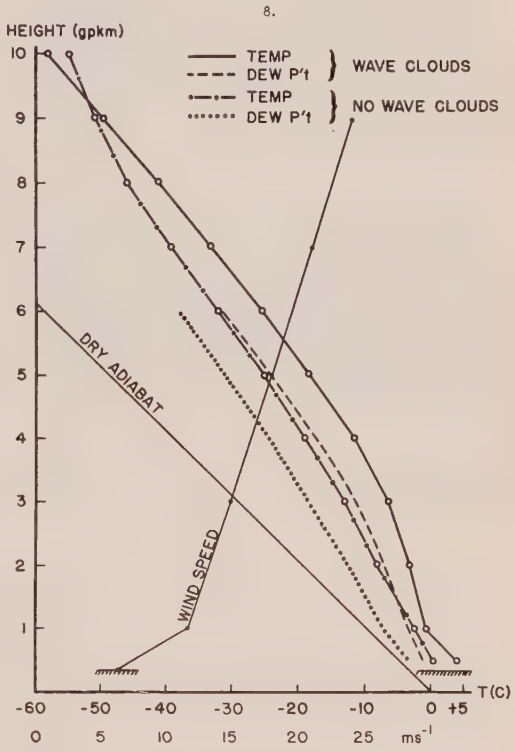


Fig. 9. Mean sounding curves for the radiosonde station at Frösön for March 1953. The soundings (altogether 19) made when wave clouds have occurred, have been compared to those (43) when no wave clouds occurred. The diagram also shows the mean wind speed for different levels, estimated from 50 cases with wave clouds (after LARSSON, 1954).

able higher degree of stability was found in the layer 1—3 km and a somewhat smaller stability above 4 km than was the case when no wave clouds were observed. In the same figure is also shown the average wind speed from 50 cases when lee wave clouds were observed. In most cases the wind speed increased with height but there were many cases in which the wind speed was constant, except in the friction layer. In no case were lee wave clouds observed when the velocity decreased with height.

We observe that the temperature distribution in the case May 9, 1955 is very similar to the one in fig. 9 which represents the mean sounding when no wave clouds occurred. This seems contradictory because waves certainly existed this day. Even with a fairly large change in the stability we should have obtained waves with sufficiently strong amplitude to give rise to wave clouds. However, it is not unreasonable to expect that this type of stability distribution with height is likely to occur when the flow is non-laminar. Absence of lee waves may thus be attributed to the fact that the flow is non-laminar and not to the direct effect of the temperature distribution.

5. Comparison with the "parcel method"

It may be of interest to compare the present method with the very most simple one which is derived by considering the vertical oscillation of an air parcel in the atmosphere. The well-known formula for the wavelength obtained in this way reads

$$L^* = 2\pi U \sqrt{\frac{T}{g(\gamma_a - \gamma)}} \tag{31}$$

where the symbols have the same meaning as before.

Making use of the idealized distributions of the wind and stability the wavelength L^* has been evaluated as a function of height. The result of these computations is shown in the figures 4—7. In all four cases the wavelength increased with height which certainly is not the case in reality. Judging from the figures it is quite clear that the wavelength obtained with this method is connected with the shortest wavelength given by the solution to the wave equation $L^* = L_1$ at about 3 km height. In the two first cases the second wave solution is the predominant one and consequently the two methods differ quite strikingly. In the last two cases, however, the "parcel method" could have been used with some success, provided we had known the height at which the formula should be applied or if we had used representative mean values of the wind and stability with respect to height.

Acknowledgement

The author is greatly obliged to Mr. L. Larsson (Royal Swedish Air Force) for his kindness to lend me his material of observations of lee wave clouds. The author also wishes to thank Dr. P. Welander for very constructive comments and discussions.

REFERENCES

Boole, G., 1844: On a general method in analysis. *Phil. Trans. 1844*, **11**, pp. 225—282.

Colson, De Ver, 1952: Result of Double Theodolite Observations at Bishop, Cal., in connection with the "Bishop Wave" Phenomena. *Bull. Amer. Met. Soc.*, **33**, No. 3, pp. 107—116.

Coulomb, M J., 1936: Sur les zéroes des fonctions de Bessel considérés comme fonction de l'ordre. *Bull. Sci. Math., 2:e Ser.*, **60: 1**, Paris pp. 297—302.

Döös, B. R., 1958: *The effect of the vertical variation of wind and stability on mountain waves.* Technical Report No. 13, Contract No. Nonr-1600(00). Office Naval Research.

Eliassen, A. and Palm, E., 1954: *Energy flux for combined gravitational sound waves.* Institute for Weather and Climate Research, Publication No.1.

Foldvik, A. and Palm, E., 1959: *On two-dimensional waves with special emphasis on longer waves.* Institute for weather and climate research, Oslo, Report No. 4.

Förchtgott, J., 1949: Wave streaming in the lee of mountain ridges. *Bull. Met. Czech.*, Prague, **3**, p. 49.

Höiland, E., 1951: *Fluid flow over a corrugated bed.* Appendix A, Fifth Progress Report, Contract No. AF 19(122)-263. Air Force Cambridge Research Center, Cambridge, Massachusetts.

Jahnke, E. and Emde, F., 1948: *Tables of functions with formulae and curves.* Fourth edition, Dover Publications, New York.

Kelvin, Lord, 1886: On stationary waves in flowing water. *Phil. Mag.* **5: 22**, pp. 353—357, 445—452, 517—530.

Larsson, L., 1954: Observations of lee wave clouds in the Jämtland mountains, Sweden. *Tellus* **6**, pp. 124—138.

—, 1958: *A method to study the structure of standing lee waves and rotors.* 7th OSTIV congress, Leszno, Poland.

Lyra, G., 1943: Theorie der stationären Leewellenströmung in freier Atmosphäre. *Z. angew. Math. Mech.* **26**, pp. 1—28.

Palm, E., 1953: On the formation of surface waves in a fluid flowing over a corrugated bed and on the development of mountain waves. *Astrophy. Norveg.*, **5: 3**, 130 pp.

Palm, E. and Foldvik, A., 1960: Contribution to the theory of twodimensional mountain waves. *Geof. Publ 21.* No. 6.

Queney, P., 1947: *Theory of perturbations in stratified currents with application to airflow over mountain bariers.* Dept. of Meteor., Univ. of Chicago, Misc. Rep. No. 23.

— 1948: The problem of air flow over mountains: A summary of theoretical studies. *Bull. Amer. Met. Soc.*, **29**, pp. 16—26.

Rayleigh, Lord, 1883: The form of standing waves on the surface of running water. *Prod. Lond. Math. Soc.*, **15**, pp. 69—78.

Sawyer, J. S., 1960: Numerical calculation of the displacements of a stratified airstream crossing a ridge of small height. *Quart. J. R. Met. Soc.*, **86**, pp. 326—345.

Scorer, R. S., 1949: Theory of waves in the lee of mountains. *Quart. J. R. Met. Soc.*, **75**, pp. 41—56.

Wurtele, M. G., 1953: *The initial value lee wave problem in an isothermal atmosphere.* Scientific Report No. 3, Sierra Wave Project, Contract No. AF (122)-263, Air Force Cambridge Research Center, Cambridge, Mass.

— 1955: The transient development of a lee wave. *J. cf Marine Res.*, **14: 1**, pp. 1—13.

Zierep, J., 1957: Neue Forschungsergebnisse aus dem Gebiet der atmosphärischen Hinderniswellen. *Beitr. Phys. Atmos.* **29: 3**, pp. 143—153

A Preliminary Study of the Dynamics of Transient, Planetary Waves in the Atmosphere

By A. WIIN-NIELSEN,[1] Joint Numerical Weather Prediction Unit, Air Weather Service

(Manuscript received March 24, revised version May 29, 1961)

Abstract

Recent studies of the planetary waves in the atmosphere have suggested that the vorticity equation assumes a quasi-stationary character for motions whose scale is comparable to the radius of the earth. These tentative results are adopted as a working hypothesis, and a preliminary study is made of the stability and structure of the waves which are possible in an atmospheric model which is formulated on the basis of the working hypothesis.

It is shown that the assumed stationary form of the vorticity equation acts as a filter which eliminates the Rossby type waves from the solution of the linearized equations in the two-parameter case. The wave-solutions in a linearized form of a three-parameter model are next found and investigated. The major results are that only slowly moving waves are possible and that unstable wave-solutions will appear for sufficiently large vertical wind-shears. A modified form of the model in which horizontal advection of vorticity is also included is analyzed, and it is shown that a wave-length of maximum instability occurs.

A prediction model which can be used in studies of the planetary waves is formulated. Since the model will apply to planetary scales only, it can be used to study the importance of non-linear interactions between the planetary scales and smaller scales.

1. Introduction

The studies of the planetary motion, by which we here mean the motion whose horizontal scale is of the same order of magnitude as the radius of the earth, have in the past mainly been confined to the stationary state. Well known examples are the investigations by CHARNEY and ELIASSEN (1949) of the importance of the large-scale mountains and of friction for the existence of the planetary waves and SMAGORINSKY'S (1953) study of the planetary waves created by large-scale heat sources and sinks and modified by friction.

[1] Present affiliation: National Center for Atmospheric Research, Boulder, Colorado.

Descriptive studies of the motion and amplitude of the planetary waves contained in the papers by ELIASEN (1958) and MARTIN (1958) show that these waves are slowly moving systems which oscillate around certain preferred geographical positions, and that the amplitude changes in time in such a way that it may double in a few days. Any theoretical model which can be applied successfully to the planetary scale must at least possess these features.

The application of the quasi-geostrophic theory in the adiabatic and frictionless formulation to the largest scales of atmospheric motions has shown that this theory predicts a

rather rapid retrogression of the planetary waves. This result follows from the linear perturbation theory and has been verified also in the non-linear case by actual time-integrations, at least in the one- and two-parameter case. In contrast to these predictions the planetary waves are observed to move slowly toward east or west.

It has recently been pointed out by BURGER (1958) from scale considerations that the β-term and that part of the divergence term which contain the Coriolis' parameter are the largest terms in the vorticity equation when the wave-length of the disturbance is comparable with the radius of the earth. In view of this result one may state that any model which include only one of the two terms is inconsistent with respect to the motion on the planetary scale. The non-divergent, one-parameter model as well as the most simple two-parameter model are both inconsistent from this point of view, because the first model includes only the β-effect, and the estimate of divergence inherent in the second model becomes very small when the scale becomes large and does not balance the β-term.

The different modified forms of the one- and two-parameter models which have been formulated (CRESSMAN 1958, WIIN-NIELSEN 1959 and 1961 b) try to correct this inconsistency by retaining the β-term and give a possible estimate of the divergence based upon the observed structure of the atmosphere. These corrected forms of the prognostic equations are therefore of a semi-empirical nature.

The simplest model within the hierarchy of quasi-geostrophic models which may differ considerably in its predictions on the planetary scale is therefore the three-parameter model. A diagnostic analysis of the distribution of divergence and vertical velocity in this model (WIIN-NIELSEN, 1961a) shows that the important factors determining the divergence are the vertical variations of the static stability and the horizontal wind, especially the curvature of the vertical profile of the horizontal wind. Since the divergence in the mid-troposphere according to the analysis depends on the deviation from linearity of the wind-profile, it is evident that this factor has not been considered in the earlier perturbation analysis (CHARNEY, 1947). It is still too early to say whether the forecasts made with this model will produce reasonably accurate forecasts on the planetary scale, but later tests of the model will give a firm answer to this question. It is, however, interesting to note that forecasts made recently using the primitive equations in a five-level representation gave no rapid retrogression of the planetary waves (HINKELMANN, results presented at the International Symposium on Numerical Weather Prediction, Tokyo, 1960).

Even if we therefore approach the state where we are able to integrate more complete forms of the prognostic equations than those based upon the quasi-geostrophic theory it is only through a systematic use of imperfect equations that we can gain further understanding of the phenomena we observe as LORENZ (1960) recently has stressed.

WELANDER (1960) has recently extended BURGER's (1958) study by deriving a general wave-equation for frictionless and adiabatic flow. He solved the wave-equation by specializing to the case of very long waves, to a linear wind profile and a constant static stability and by considering only the leading terms in the series for the coefficients in the wave-equation. The solution shows in this first order approximation that the waves will move slowly with a speed which is independent of the wave-number, but depends on the product of the Rossby and Richardson numbers.

The present paper may be considered as an extension of Welander's study. As will be shown later it turns out that the special case considered by Welander is identical to the use of the geostrophic assumption not only for advection purposes but also for estimates of the divergence, or in other words a balance in the vorticity equation between the β-term and $f\nabla \cdot \mathbf{v}$. It is the purpose of this study to investigate this assumption in the more general case where we allow deviations from a linear wind profile, a vertical variation of the static stability and also the presence of unstable solutions. The general wave equation is extremely difficult to solve in the general case, and we shall therefore restrict the investigation to a vertical resolution corresponding to a three-parameter model, which is the smallest resolution which allows the possibility of unstable solutions.

In the following we shall in most cases neglect the effects of mountains, friction, and heat sources. The reason for this is not that these effects are considered to be of no importance,

but the neglect is made partly for convenience and partly because we want to isolate the internal dynamical factors in the system.

2. A filter for waves of the Rossby-type

The result of the scale analysis of the vorticity equation when applied to the planetary scale is according to BURGER (1958) that the equation takes the approximate form:

$$\mathbf{v} \cdot \nabla f + f \nabla \cdot \mathbf{v} = 0. \tag{2.1}$$

Another result obtained from the scale analysis is that the geostrophic assumption is well satisfied on the planetary scale. It is further interesting to note that a vorticity equation of the same form as (2.1) would be obtained directly from the geostrophic relations.

It is the main purpose of the present study to investigate the wave disturbances which are possible in a model atmosphere in which (2.1) is used at a diagnostic relation while the prognostic equation is the thermodynamic equation. Before we investigate the stability properties and structure of such disturbances we shall in this section show that the application of a vorticity equation in which the local time-derivative of vorticity is neglected formally acts as a filter which eliminates the Rossby type waves as solutions to the linearized equations. We shall illustrate this formal argument by comparing the well-known two-parameter model used in short-range numerical prediction with the corresponding model, in which the time-derivative of the vorticity is neglected. The basic equation for the latter model may be written with $\omega = 0$ for $p = 0$ and $p = 100$ cb.

$$\begin{aligned} &\mathbf{v}_1 \cdot \nabla \zeta_1 + \beta v_1 = -\frac{f_0}{P}\,\omega_2 \\ &\mathbf{v}_3 \cdot \nabla \zeta_3 + \beta v_3 = -\frac{f_0}{P}\,\omega_2 \\ &\frac{\partial \varphi'}{\partial t} + \mathbf{v}_2 \cdot \nabla \varphi' - \frac{P}{2}\,\sigma_2\,\omega_2 = 0, \end{aligned} \tag{2.2}$$

where the subscripts 1, 2, and 3 refer to the 25, 50, and 75 cb levels, respectively, and where $P = 50$ cb, $\varphi' = \frac{1}{2}(\varphi_1 - \varphi_3)$ and $\mathbf{v}_2 = \frac{1}{2}(\mathbf{v}_1 + \mathbf{v}_3)$. $\sigma = -\alpha\,\partial \ln \Theta / \partial p$ is the measure of static stability.

It will be seen that if the terms $\partial \zeta_1/\partial t$ and $\partial \zeta_3/\partial t$ were added on the left sides of the first two equations in (2.2), we would have the conventional two-parameter model.

It is convenient to add and subtract the first two equations in the system (2.2). We obtain

$$\begin{aligned} &\mathbf{v}_2 \cdot \nabla \zeta_2 + \mathbf{v}' \cdot \nabla \zeta' + \beta v_2 = 0 \\ &\mathbf{v}_2 \cdot \nabla \zeta' + \mathbf{v}' \cdot \nabla \zeta_2 + \beta v' = \frac{f_0}{P}\,\omega_2 \end{aligned} \tag{2.3}$$

Combining the last of the equations (2.3) with the thermodynamic equation we finally have the system

$$\begin{aligned} &\mathbf{v}_2 \cdot \nabla \zeta_2 + \beta v_2 + \mathbf{v}' \cdot \nabla \zeta' = 0 \\ &\frac{\partial \varphi'}{\partial t} + \mathbf{v}_2 \cdot \nabla \varphi' - \frac{\sigma P^2}{2 f_0}\,(\mathbf{v}_2 \cdot \nabla \zeta' + \mathbf{v}' \cdot \nabla \zeta_2) - \\ &\qquad - \frac{\sigma P^2}{2 f_0}\,\beta v' = 0. \end{aligned} \tag{2.4}$$

It will be seen from (2.4) that in spite of the neglect of the local change of vorticity we still have a complete set of equations if the winds and vorticities are approximated geostrophically.

Linearizing the equations (2.4) and assuming the perturbations to be independent of the meridional direction we obtain with $q = 2 f_0^2/\sigma P^2$:

$$\begin{aligned} &\frac{\partial \varphi'}{\partial t} + U_2 \frac{\partial \varphi'}{\partial x} - U' \frac{\partial \varphi_2}{\partial x} - \\ &- \frac{1}{q}\left(U_2 \frac{\partial^3 \varphi'}{\partial x^3} + U' \frac{\partial^3 \varphi_2}{\partial x^3} \right) - \frac{\beta}{q} \frac{\partial \varphi'}{\partial x} = 0 \\ &U_2 \frac{\partial^3 \varphi_2}{\partial x^3} + U' \frac{\partial^3 \varphi'}{\partial x^3} + \beta \frac{\partial \varphi_2}{\partial x} = 0 \end{aligned} \tag{2.5}$$

Wave solutions of the type

$$\varphi_2 = A_2\, e^{ik\,(x - ct)},\ \varphi' = A'\, e^{ik\,(x - ct)} \tag{2.6}$$

can be found by substituting (2.6) in (2.5). The condition that the determinant has to vanish in the resulting two homogeneous, linear equations in order to obtain non-trivial solution leads to the wave-speed formula:

$$c = \left(1 + \frac{k^2}{q}\right) U_2 - \frac{\beta}{q} = \left(1 - \frac{k^2}{q}\right) \frac{U' k^2}{k^2 U_2^2 - \beta} \tag{2.7}$$

We notice first of all that we get only one root in the frequency equation compared to

Table 1

L, km..	1,000	2,000	3,000	4,000	5,000	6,000	7,000	8,000	9,000	10,000	11,000	12,000	13,000	14,000	28,000	∞
c, m/sec	12.4	13.1	15.4	23.1	∞	− 6.9	− 0.3	2.2	3.3	4.1	4.4	4.7	5.0	5.1	5.8	6.0

two roots in the similar analysis of the usual two-parameter model. It is to be expected that the root (2.7) is closely connected with one of the wave-speeds appearing in the conventional two-parameter model. These are given for instance in a paper by the author (1959, eq. (6.7)). For $r = 0$ in the reference formula we find that one of the speeds approaches the Rossby wave speed

$$c = U_2 - \beta/k^2 \to -\infty \text{ for } k \to 0, \tag{2.8}$$

while the other in the limit for $k \to 0$ is

$$c = U_2 - \beta/q. \tag{2.9}$$

For $k \to 0$ we find from (2.7) that

$$c = U_2 - \beta/q, \tag{2.10}$$

which means that the neglect of the local time-derivative in the vorticity equation act as a filter which eliminates the Rossby-type waves as solutions to the linearized equations in the two-parameter case. The wave-speed (2.7) is given as a function of wave-length in Table 1 for $U_2 = U' = 10$ m sec^{-1}, $q = 4 \times 10^{-12}$ m^{-2} and $\beta = 16 \times 10^{-12}$ m^{-1} sec^{-1}. It is seen from the table, as can also be seen directly from (2.7), that c approaches infinity for the stationary Rossby wave and that fast-moving waves also may appear for small wave-lengths if $U_2 \neq U'$.

If we take the further step to disregard the advection of relative vorticity in (2.2) we get with the same notations as before the following set of equations corresponding to (2.5)

$$\begin{aligned} \frac{\partial \varphi'}{\partial t} + U_2 \frac{\partial \varphi'}{\partial x} - \frac{\beta}{q} \frac{\partial \varphi'}{\partial x} &= 0 \\ \frac{\partial q_2}{\partial x} &= 0 \end{aligned} \tag{2.11}$$

With the same type of perturbations we obtain now instead of (2.7) the wave-speed

$$c = U_2 - \frac{\beta}{q}, \tag{2.12}$$

which shows that if we in accordance with Burger assume the form (2.1) of the vorticity equation we obtain the result from a linear analysis that the wave-speed becomes independent of the wave-number, but naturally agrees with the wave-speed (2.7) for small values of k, which expresses the fact that the vorticity advection is small compared to the β-term and $f \nabla \cdot \mathbf{v}$ for very long waves.

The last result throws also some light on the results obtained by Welander (1960), because it is possible to show that the final equation solved by him and obtained considering only the leading terms in a more general equation also can be obtained from the equations:

$$\begin{aligned} u &= -\frac{1}{f} \frac{\partial \varphi}{\partial y} \\ v &= +\frac{1}{f} \frac{\partial \varphi}{\partial x} \\ \partial u/\partial x + \partial v/\partial y &- \frac{\tan \gamma}{R} v + \frac{\partial \omega}{\partial p} = 0 \\ \frac{\partial}{\partial t}\left(\frac{\partial \varphi}{\partial p}\right) &+ \mathbf{v} \cdot \nabla \left(\frac{\partial \varphi}{\partial p}\right) + \sigma \omega = 0, \end{aligned} \tag{2.13}$$

of which the first two simply express the geostrophic wind-law, the third is the continuity equation, and the last equation is the thermodynamic equation in the adiabatic case. x and y are now the quasi-coordinates used by Welander.

Assuming as Welander does that σ is a function of pressure only, we obtain from the last equation:

$$\begin{aligned} -\frac{\partial \omega}{\partial p} = \frac{\partial}{\partial t}\left(\frac{\partial}{\partial p}\left(\frac{1}{\sigma}\frac{\partial \varphi}{\partial p}\right)\right) + \mathbf{v} \cdot \nabla \frac{\partial}{\partial p}\left(\frac{1}{\sigma}\frac{\partial \varphi}{\partial p}\right) + \\ + \frac{\partial \mathbf{v}}{\partial p} \cdot \nabla \left(\frac{1}{\sigma}\frac{\partial \varphi}{\partial p}\right) \end{aligned} \tag{2.14}$$

The last term in (2.14) disappears due to the geostrophic assumption, while the left side is evaluated from the first three equations in (2.13) from which we obtain:

$$\frac{\partial \omega}{\partial p} = \frac{\beta}{f^2} \frac{\partial \varphi}{\partial x} \tag{2.15}$$

When we combine (2.14) and (2.15) we obtain a single equation in the geopotential. We introduce disturbances of the form $\varphi = A(p)e^{ik(x-ct)}$ on a zonal, geostrophic current, which means that

$$f\frac{\partial U}{\partial p} = -\frac{\partial^2 \Phi}{\partial y \partial p}. \tag{2.16}$$

The resulting wave-equation is

$$(U-c)\frac{d}{dp}\left(\frac{1}{\sigma}\frac{dA}{dp}\right) + \left(\frac{\beta}{f^2} - \frac{d}{dp}\left(\frac{1}{\sigma}\frac{dU}{dp}\right)\right)A = 0 \tag{2.17}$$

which agrees with Welander's final equation when his scaling is introduced in (2.17).

We have therefore shown that the problem solved by Welander is equivalent to solving the perturbation problem in a model atmosphere governed by the set of equations (2.13). Reduced to a two-parameter model it is equivalent to the set of equations (2.11) from which we have derived the result that the Rossby-type waves are filtered out and the remaining waves move with a speed independent of the wave-number, a result also obtained by Welander.

The development which has been given so far in the present paper is naturally of a formal nature. Nothing has been presented which shows that the wave types remaining after the Rossby-waves are filtered out describes the motion of the very long waves in the atmosphere. However, the two-parameter model which has been analyzed in the present section is in agreement with the findings reported by Burger (1958) and Welander (1960). We shall in the following section make a further study of the stability and structure of planetary waves based upon the working hypothesis that a neglect of the local time derivative serves to filter the Rossby-type waves, and that the remaining wave solutions describe the internal dynamics of the planetary waves in the atmosphere. The general wave-equation derived by Welander is extremely difficult to solve. We shall therefore first investigate a simple three-parameter model based on a balance between the β-term and the divergence term in the vorticity equation. This model will in a later section be generalized to include also the advection of relative vorticity.

3. A simple three-parameter mode

In this section we shall analyze a three-parameter model where we, for simplicity, shall assume the following basic equations, which are in agreement with the scale analysis for the planetary waves:

$$\beta v = f_0 \frac{\partial \omega}{\partial p}$$
$$\frac{\partial}{\partial t}\left(\frac{\partial \varphi}{\partial p}\right) + \mathbf{v}\cdot\nabla\left(\frac{\partial \varphi}{\partial p}\right) + \sigma\omega = 0. \tag{3.1}$$

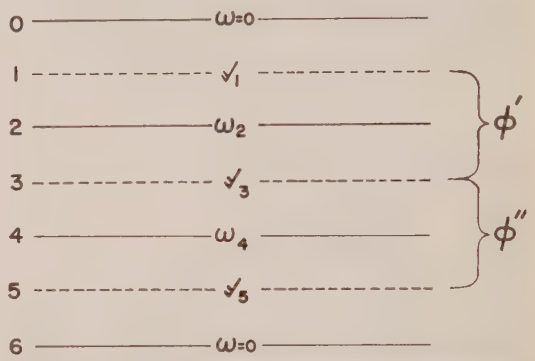


Figure 1. A schematic representation of the theoretical model. ω_2 and ω_4 are the internal vertical velocities, $\mathbf{v}_1$, $\mathbf{v}_3$ and $\mathbf{v}_5$ the horizontal winds and φ' and φ'' the thicknesses over the indicated layers.

The notations which will be used in the following are entered on figure 1. We shall in order to isolate the internal dynamical factors assume that $\omega_6 = 0$. The equations for the model are then

$$\beta v_1 = \frac{f_0}{P}\omega_2, \quad \beta v_3 = \frac{f_0}{P}(\omega_4 - \omega_2), \quad \beta v_5 = -\frac{f_0}{P}\omega_4$$
$$\frac{\partial \varphi'}{\partial t} + \mathbf{v}_3\cdot\nabla\varphi' - P\sigma_2\omega_2 = 0 \tag{3.2}$$
$$\frac{\partial \varphi''}{\partial t} + \mathbf{v}_3\cdot\nabla\varphi'' - P\sigma_4\omega_4 = 0.$$

Adding the three first equations in (3.2) we obtain

$$v_1 + v_3 + v_5 = 0. \tag{3.3}$$

We introduce $v_1 = v_3 + v'$ and $v_5 = v_3 - v''$ and may then write (3.3) in the form:

$$v_3 = \frac{-1}{3}\cdot(v' - v'').$$

The last two equations in (3.2) are next written in a linearized form after we have substituted for ω_2, ω_4, and v_3. After rearrangement the equations are:

$$\frac{\partial \varphi'}{\partial t} + \left(U_3 + \frac{1}{3}U' - \frac{2}{3}\frac{\beta}{q_2}\right)\frac{\partial \varphi'}{\partial x} - \frac{1}{3}\left(U' + \frac{\beta}{q_2}\right)\frac{\partial \varphi''}{\partial x} = 0$$

$$\frac{\partial \varphi''}{\partial t} + \left(U_3 - \frac{1}{3}U'' - \frac{2}{3}\frac{\beta}{q_4}\right)\frac{\partial \varphi''}{\partial x} + \frac{1}{3}\left(U'' - \frac{\beta}{q_4}\right)\frac{\partial \varphi'}{\partial x} = 0 \qquad (3.5)$$

where $q_2 = f_0^2/\sigma_2 P^2$ and $q_4 = f_0^2/\sigma_4 P^2$.

Assuming next perturbations of the form

$$\varphi' = A' e^{ik(x-ct)}$$
$$\varphi'' = A'' e^{ik(x-ct)} \qquad (3.6)$$

we can by substituting (3.6) in (3.5) derive a frequency equation whose solution may be written

$$c = U_3 + 1/6\,(U' - U'') - \beta/3\,(1/q_2 + 1/q_4) \pm \frac{1}{2}\sqrt{D}. \qquad (3.7)$$

with

$$D = \frac{1}{9}\Big[(U' - U'')^2 + 4\beta^2(1/q_2^2 + 1/q_4^2 - 1/q_2 q_4) + 4\beta[(U'/q_4 - U''/q_2) + (U' + U'')(1/q_4 - 1/q_2)]\Big] \qquad (3.8)$$

We notice in the solutions (3.7) that the wave-speed is independent of the wave-number, which verifies that the Rossby-type waves are filtered out of the system. Before we investigate the general case we may consider some special cases.

Let us first suppose that $U' = U'' = U_T$ and $q_2 = q_4 = q_m$. The three-parameter model has in this case the same structure as a two parameter model described by the set of equations (2.11). We find the two wave-speeds to be

$$C_1 = U_3 - \frac{1}{3}\frac{\beta}{q_m}, \quad C_2 = U_3 - \frac{\beta}{q_m}, \qquad (3.9)$$

which are comparable to (2.12). If therefore the wind increases linearly with decreasing pressure and the stability is quasi-constant, we find only stable, slowly moving waves as solutions.

In the next case we shall assume that $U' = U'' = U_T$, but $q_2 < q_4$, which corresponds to an increasing stability with decreasing pressure. In this case we find

$$D = \frac{4}{9}\beta\,[\beta\,(1/q_2^2 + 1/q_4^2 - 1/q_2 q_4) - 3\,U_T\,(1/q_2 - 1/q_4)] \qquad (3.10)$$

It follows from our assumption that $q_2^{-1} - q_4^{-1} > 0$. It is further easy to show that $q_2^{-1} + q_4^{-1} - q_2^{-1} q_4^{-1} > 0$ for all values of q_2 and q_4. (3.10) shows therefore that $D < 0$ for sufficiently large values of U_T, and that the waves are unstable. The critical value of U_T denoted U_{TC}, corresponding to neutral stability is

$$U_{TC} = \frac{\beta}{3}\left(\frac{1}{q_2} - \frac{1}{q_4} - \frac{1}{q_4 - q_2}\right). \qquad (3.11)$$

Introducing $r = \sigma_2/\sigma_4$ we find for $\beta = 16 \times 10^{-12}$ m^{-1} sec^{-1}, $f_0 = 10^{-4}$ sec^{-1} and $\sigma_4 = 1.7$ MTS-units that

$$U_{TC} = \frac{r^2 - r + 1}{r - 1}. \qquad (3.12)$$

Table 2 shows corresponding values of r and (dU/dz) computed from (3.12). It is seen that the waves will be unstable for rather small values of dU/dz over a wide range of values of r. It is interesting to note that the smallest value of U_{TC} is found for $r = 2$.

The previous result shows that since the (complex) wave-speed is independent of the wave-number, we will find that if the wave is unstable, instability will occur for all wave-lengths. The amplification rate will depend upon the wave-length even if the wave-speed is independent of wave-number, as we can see from the following expression for the time it will take to double the amplitude:

$$T = \frac{\ln 2}{kC_i} \qquad (3.13)$$

Table 2

$r = \sigma_2/\sigma_4$	1.1	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	10.0
dU/dz, m sec^{-1} km^{-1}	1.67	0.53	0.45	0.48	0.53	0.59	0.65	0.72	0.80	1.52

Table 3

N	1	2	3	4	
dU/dz, m sec^{-1} km^{-1}	65.0	16.7	7.3	4.7	1 day
dU/dz, m sec^{-1} km^{-1}	16.3	4.2	1.8	1.2	2 days

where C_i is the imaginary part of the root. Table 3 contains the values of dU/dz in m sec^{-1} km^{-1} necessary to double the amplitude in 1 and 2 days for different values of N, the zonal wave-number.

It is seen from Table 3 that the instability is rather weak for the small wave-numbers. Normal values of dU/dz are at a maximum in winter and may be as large as 4—5 m sec^{-1} km^{-1}. If the present case were characteristic of the atmosphere we would therefore find a doubling time of the order of two to three days for wave-numbers, $N = 1$ to 3. However, the instability is probably less in the atmosphere than shown in the present example as will be seen from the treatment of the more general case.

We notice further from (3.13) that the amplification time will be extremely short for large values of k. This is of no meteorological importance since we know that the present model does not apply to medium and small scale motion in the atmosphere. If, however, a numerical integration were to be made on the basis of the equations (3.2) a serious problem would appear in controlling the small scale motion which inevitably would appear in the course of the integration.

We shall finally shortly discuss the more general case in which $U' \neq U''$ and $q_2 \neq q_4$. It is convenient to introduce the notations

$$U_T = \frac{1}{2}(U' + U''), \quad \Delta U = \frac{1}{2}(U' - U''). \quad (3.14)$$

U_T is a measure of the mean shear, while ΔU is a first approximation to a measure of the curvature of the vertical profile of the horizontal wind as can be seen from the expression:

$$\frac{d^2U}{dp^2} \approx \frac{U_1 + U_5 - 2U_3}{P^2} = \frac{2}{P^2}\Delta U.$$

In the middle latitudes we find in general $\Delta U > 0$ as can be seen from wind statistics (WIIN-NIELSEN, 1961a).

The expression for D becomes now with the notations (3.14):

$$D = \frac{4}{9}[\beta^2(q_2^{-2} + q_4^{-2} - q_2^{-1}q_4^{-1}) - \quad (3.16)$$
$$- 3\beta U_T(q_2^{-1} - q_4^{-1}) + \beta\Delta U(q_2^{-1} + q_4^{-1}) + \Delta U^2]$$

Comparing (3.16) and (3.10) it is seen that positive values of ΔU will increase the values of D and therefore stabilize the waves. The critical values of dU/dz for different values of ΔU are given in Table 4 assuming a standard ratio of $r = 4$. The stabilizing influence of positive values of ΔU is clearly seen in Table 4, but it is also evident that the planetary waves would be unstable in many actual situations for moderately large values of dU/dz.

4. On the structure of the unstable waves

The unstable waves which are possible solutions to the linearized equations in the present model should have a vertical structure which is in a gross sense comparable to the structure of the planetary waves in the atmosphere if our model is at all an approximation to reality. In this section we will derive certain properties of the unstable waves. We shall especially be interested in the vertical slope of the waves and the transport of sensible heat caused by the unstable waves. Due to the fact that we, in the previous section, have assumed that no horizontal shear exist in the basic current and that the

Table 4

ΔU, m sec^{-1}	− 20	− 15	− 10	− 5	0	5	10	15	20
dU/dz, m sec^{-1} km^{-1}	1.19	0.65	0.36	0.38	0.65	1.22	2.04	3.17	4.55

waves are independent of the meridional direction we obtain no meridional convergence of momentum or heat transport. It is, however, possible to evaluate the transport of sensible heat, while the momentum transport vanishes.

The general solution in the three-parameter case was obtained from the equations (3.2) in a linearized form. When we substitute the solution for the wave speed back into the original linearized equations we obtain two independent relations between A' and A_3 and between A'' and A_3. These equations can be written in the form:

$$\begin{aligned}(C - U_3 + \beta/q_2)\, A' &= -(U' + \beta/q_2)\, A_3 \\ (C - U_3 + \beta/q_4)\, A'' &= -(U'' - \beta/q_4)\, A_3\end{aligned} \tag{4.1}$$

where C is determined by (3.7) and (3.8). In the following we shall concentrate on the unstable wave solutions. The wave-speed C can for these solutions be written in the form:

$$C = C_r + iC_i,\ C_i > 0. \tag{4.2}$$

It is seen from (4.1) that the amplitudes A', A'', and A_3 in general are complex numbers. If we arbitrarily set the amplitude at level 3 to unity, i.e., $A_3 = 1$, we can find the amplitudes A' and A'' from (4.1). Denoting

$$A' = A'_r + iA_i,\ A'' = A''_r + iA''_i \tag{4.3}$$

we can, by substitution in (4.1) and separating the real and imaginary parts, solve for the four quantities, A'_r, A'_i, A''_r, and A''_r. The result is

$$\begin{aligned}A'_r &= -\frac{(U' + \beta/q_2)\,(C_r - U_3 + \beta/q_2)}{(C_r - U_3 + \beta/q_2)^2 + C_i^2} \\ A'_i &= +\frac{(U' + \beta/q_2)\, C_i}{(C_r - U_3 + \beta/q_2)^2 + C_i^2} \\ A''_r &= -\frac{(U'' - \beta/q_4)\,(C_r - U_3 + \beta/q_4)}{(C_r - U_3 + \beta/q_4)^2 + C_i^2} \\ A''_i &= +\frac{(U'' - \beta/q_4)\, C_i}{(C_r - U_3 + \beta/q_4)^2 + C_i^2}\end{aligned} \tag{4.4}$$

These expressions are most easily evaluated in the case where $U' = U''$ and $q_2 < q_4$. In this case we have:

$$C_r - U_3 = -\frac{\beta}{3}\left(\frac{1}{q_2} + \frac{1}{q_4}\right),\ C_i = \frac{1}{3} \times \tag{4.5}$$
$$\times \left[3U_T\left(\frac{\beta}{q_2} - \frac{\beta}{q_4}\right) - \left(\frac{\beta^2}{q_2^2} + \frac{\beta^2}{q_4^2} - \frac{\beta^2}{q_2 q_4}\right)\right]^{\frac{1}{2}}$$

We notice first that if $U_T \leq \beta/q_4$ we have only stable solutions. In the case we are discussing we can therefore safely assume that $U_T > \beta/q_4$. This means that A'_i and A''_i are positive numbers. It is obvious that A'_r is negative, while A''_r for standard values of q_2 and q_4 turn out to be positive.

The values of the relative amplitudes A'_r, A'_i, A''_r, and A''_i were computed for two different cases. The values of q_2 and q_4 were computed for 45°N from average data for the winter and summer 1950 (Peixoto, 1960). The difference between the two seasons was so small for this latitude that the same values were accepted for both seasons, $\sigma_2 = 7.0$ MTS-units and $\sigma_4 = 1.8$ MTS-units. With these values we find for 45° N $\beta/q_2 = 12$ m sec^{-1} and $\beta/q_4 = 3$ m sec^{-1}. We find further that

$$U_T \approx P\frac{dU}{dp} = 5.0\,\frac{dU}{dz}, \tag{4.6}$$

where dU/dz is measured in m sec^{-1} km^{-1}. Mean values of dU/dz have been computed from mean meridional cross-sections for winter and summer by the author (1959) as a function of latitude. The relative amplitudes were computed for the values $dU/dz = 3$ m sec^{-1} km^{-1} and 1.5 m sec^{-1} km^{-1}.

A component of the wave solution may be written in the form:

$$\begin{aligned}\varphi' &= (A'_r + iA'_i)\, e^{kC_i t}\,(\cos k\,(x - C_r t) + \\ &\quad + i \sin k\,(x - C_r t)) \\ \varphi'' &= (A''_r + iA''_i)\, e^{kC_i t}\,(\cos k\,(x - C_r t) + \\ &\quad + i \sin k\,(x - C_r t))\end{aligned} \tag{4.7}$$

where the perturbations as usual are to be identified with the real parts of (4.7). We find therefore

$$\begin{aligned}\varphi'' &= A''_r \cos X - A''_i \sin X \\ \varphi' &= A'_r \cos X - A'_i \sin X\end{aligned},\ X = k\,(x - C_r t) \tag{4.8}$$

from which it follows that

$$\begin{aligned}\varphi_1 &= \varphi_3 + \varphi' = (1 + A'_r)\cos X - A'_i \sin X \\ \varphi_5 &= \varphi_3 + \varphi'' = (1 - A''_r)\cos X + A'' \sin X\end{aligned} \tag{4.9}$$

Any expression of the form (4.8) or (4.9) can be written in the more convenient form

$$\varphi = A \cos (X + \delta), \tag{4.10}$$

where A is the amplitude and δ is the phase.

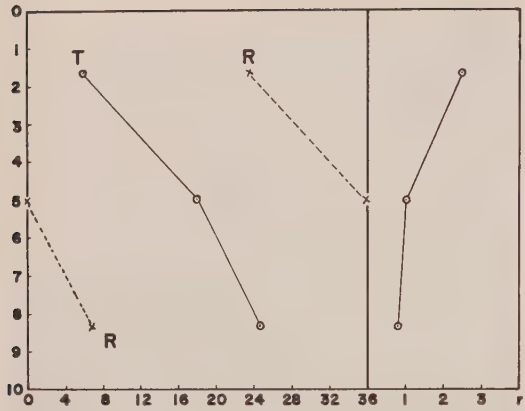


Figure 2. The vertical slope of troughs and ridges and the relative variation of amplitudes in the unstable baroclinic wave independent of wave-number for $dU/dz = 3$ m sec^{-1} km^{-1}.

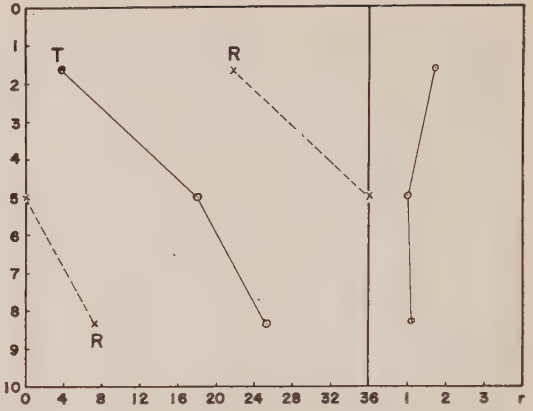


Figure 3. As figure 2, but with $dU/dz = 1.5$ m sec^{-1} km^{-1}.

Table 5

	dU/dz = 3 m sec^{-1} km^{-1}		dU/dz = 1.5 m sec^{-1} km^{1-}	
	A	δ (deg)	A	δ (deg)
φ_1	2.31	+125	1.68	+142
φ_3	1.00	0	1.00	0
φ_5	1.92	−80	1.08	−72
φ'	3.00	+141	2.55	+156
φ''	2.00	+71	1.22	+57

The result of the two computations are given in Table 5.

The position of the trough and ridge at the different levels and the vertical variation of the relative amplitude are further given in figures 2 and 3. It will be seen that the systems slope toward the west with decreasing pressure in such a way that the ridge at level 1 is approximately over the trough at level 5 and vice versa. It is further seen that the smallest amplitude is found in the mid-troposphere at level 3. The diagrams for different wave-numbers will look exactly as figures 2 and 3, because the solution is independent of wave-number. We may compare with the structure of the planetary waves in the atmosphere. Figures 4, 5, and 6 contain the positions of the troughs and ridges for wave-numbers 1, 2, and 3, respectively, computed from the normal maps of January at 50° N. Some similarity is found for

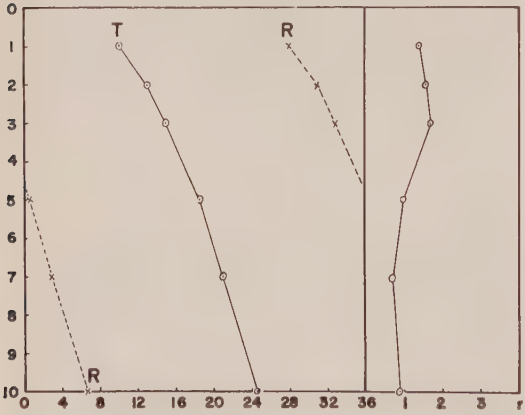


Figure 4. The vertical slope of troughs and ridges and the vertical variation of amplitude for zonal wave-number $N = 1$, computed from January normals at 50° N.

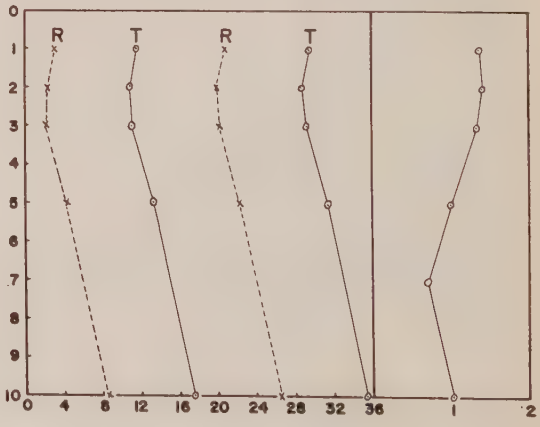


Figure 5. As figure 4, $N = 2$.

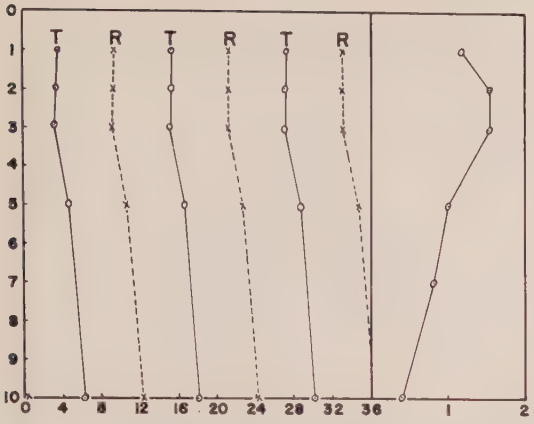


Figure 6. As figure 4, $N = 3$.

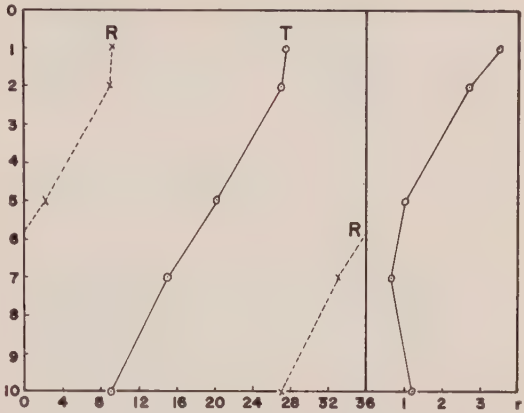


Figure 7. As figure 4, but for the July normals, $N = 1$.

wave-number 1 in the sense that the wave has approximately the same slope and also that the minimum amplitude is found in the mid-troposphere (700 mb). It is, however, evident that the wave derived from the present simplified theory slope too much in the vertical direction, especially when we compare with the observed slope of wave-numbers 2 and 3, which tend to be more vertical. Figure 7 shows in a similar way wave number 1 computed from the July normals. All the diagrams in figures 4 to 7 are extensions of the diagrams prepared by ELIASEN (1958) and can be compared with his figure 4. The significant fact is that the planetary waves have the opposite slope in summer, when computed from the normal maps. The present theory is apparently unable to produce such a slope. The reason for this is to be found among the factors which have been left out of the present model among which we may point to the meridional scale, friction, and heat-sources. The main results which we have obtained from our model is the fact that unstable solutions are possible even for the planetary scales, and we have therefore attempted to point to the instability as a factor which may account for the fact that planetary waves seem to be excited in many instances.

Some caution should naturally be taken in comparing the present theory with the slopes and amplitudes on the normal maps. In order to obtain a representative picture of the planetary waves in the different seasons one should rather obtain average slopes and amplitudes from individual maps by a Fourier analysis of the daily maps. The results of such an analysis will be reported later.

We shall finish this section by investigating the heat-transport of the unstable waves. The averaged heat-transport in the meridional direction in the upper and lower layers is proportional to

$$T_2 = \overline{\varphi' v_3}, \quad T_4 = \overline{\varphi'' v_3}, \tag{4.11}$$

where the bar means

$$\overline{(\,)} = \frac{1}{L}\int_0^L (\,)\, dx, \tag{4.12}$$

and L is the wave-length.

It is convenient to write the thickness and the geopotential at level 3 in the form (4.10), i.e.,

$$\begin{aligned} \varphi_3 &= A_3 \cos X \\ \varphi' &= A' \cos (X + \delta') \\ \varphi'' &= A'' \cos (X + \delta'') \end{aligned} \tag{4.13}$$

With these notations we find:

$$T_2 = \frac{1}{2f_0} A' A_3 \sin \delta', \quad T_4 = \frac{1}{2f_0} A'' A_3 \sin \delta''. \tag{4.14}$$

Using the values given in Table 5 it is seen that δ' and δ'' both are between 0 and 180 degrees which indicates that the unstable waves will transport sensible heat toward the north. It is interesting to note that if the average slope in the summer time, indeed, is from west to east with decreasing pressure, we

should also find a southward transport of sensible heat by the planetary waves in the summer time. Unfortunately, we do not have any information at the moment of the transport of heat in the terms of wave-numbers.

5. On the influence of advection of relative vorticity

One of the main objections against the results obtained in the two preceding sections is that instability will occur on all scales, which consequently means that the amplification rate will be extremely large for any smaller scale motion which may be present. A further result was that all scales will have the same vertical slope measured relative to the wave-length. We shall demonstrate in this section that if we include the advection of relative vorticity in the diagnostic vorticity equation in addition to the β-term and $f\nabla \cdot \mathbf{v}$, we obtain a stabilization of the shorter waves, and a wave-length of maximum instability occurs. In demonstrating this statement we shall use the same vertical resolution as previously, which is described in figure 1. We shall again for simplicity use the boundary conditions $\omega = 0$ for $p = 0$ and $p = p_6$. The system of equations for the modified model become

$$\begin{aligned}
&\mathbf{v}_1 \cdot \nabla\zeta_1 + \beta v_1 = \frac{f_0}{P}\omega_2 \\
&\mathbf{v}_3 \cdot \nabla\zeta_3 + \beta v_3 = \frac{f_0}{P}(\omega_4 - \omega_2) \\
&\mathbf{v}_5 \cdot \nabla\zeta_5 + \beta v_5 = -\frac{f_0}{P}\omega_4 \qquad (5.1) \\
&\frac{\partial\varphi'}{\partial t} + \mathbf{v}_3 \cdot \nabla\varphi' - P\sigma_2\omega_2 = 0 \\
&\frac{\partial\varphi''}{\partial t} + \mathbf{v}_3 \cdot \nabla\varphi'' - P\sigma_4\omega_4 = 0.
\end{aligned}$$

When we introduce $\mathbf{v}' = \mathbf{v}_1 - \mathbf{v}_3$, $\mathbf{v}'' = \mathbf{v}_3 - \mathbf{v}_5$, solve for ω_2 and ω_4 from the first and third equation in the system (5.1) and substitute in the remaining equations, we eliminate all vertical velocities and the system is reduced to three equations, two prognostic equations for φ' and φ'' and one diagnostic equation for φ_3. This system is:

$$\begin{aligned}
&\frac{\partial\varphi'}{\partial t} + \mathbf{v}_3 \cdot \nabla\varphi' - \frac{\sigma_2 P^2}{f_0}(\mathbf{v}_3 + \mathbf{v}') \cdot \nabla(\zeta_3 + \zeta') + \\
&\qquad + \beta(v_3 + v') = 0 \\
&\frac{\partial\varphi''}{\partial t} + \mathbf{v}_3 \cdot \nabla\varphi'' - \frac{\sigma_4 P^2}{f_0}(\mathbf{v}_3 - \mathbf{v}'') \cdot \nabla(\zeta_3 - \zeta'') + \\
&\qquad + \beta(v_3 - v'') = 0 \qquad (5.2) \\
&\mathbf{v}_3 \cdot \nabla\zeta_3 + (\mathbf{v}_3 + \mathbf{v}') \cdot \nabla(\zeta_3 + \zeta') + \\
&+ (\mathbf{v}_3 - \mathbf{v}'') \cdot \nabla(\zeta_3 - \zeta'') + \beta(3v_3 + v' - v'') = 0
\end{aligned}$$

Assuming next perturbations of the form

$$\varphi' = A'e^{ik(x-ct)},\ \varphi'' = A''e^{ik(x-ct)},\ \varphi_3 = A_3e^{ik(x-ct)}, \qquad (5.3)$$

superposed on a zonal current with no lateral shear we can replace the equations (5.2) by the following set of linear equations

$$\begin{aligned}
&\left[C - U_3 - \frac{k^2}{q_2}(U_3 + U') + \frac{\beta}{q_2}\right]A' + \\
&+ \left[U' - \frac{k^2}{q_2}(U_3 + U') + \frac{\beta}{q_2}\right]A_3 = 0 \\
&\left[C - U_3 - \frac{k^2}{q_4}(U_3 - U'') + \frac{\beta}{q_4}\right]A'' + \qquad (5.4) \\
&+ \left[U'' + \frac{k^2}{q_2}(U_3 - U'') - \frac{\beta}{q_4}\right]A_3 = 0 \\
&[\beta - k^2(U_3 + U')]A' - [\beta - k^2(U_3 - U'')]A'' + \\
&+ [\beta - k^2(3U_3 + U' - U'')]A_3 = 0.
\end{aligned}$$

The condition that the determinant has to vanish in order to give non-trivial solutions results in a quadratic equation for the phase-speed C. Due to the presence of the advection of relative vorticity the solution will now depend on the wave-number. Ascribing values to β, L, U_3, U', U'', q_2, and q_4, where L is the wave-length, it is possible to solve for the phase-speed C. Especially we can find the wave-lengths and wind-shears for which instability occurs. Constant values have been assigned to β, U_3, q_2 and q_4. By varying L and $U' = U'' = U_T$ we can compute diagrams showing the regions of instability and also the amplification rate computed from the following formula:

$$T = \frac{1}{kC_i}. \qquad (5.5)$$

T is the so-called e-folding time.

Figure 8 gives the results of such computations. The abscissa is the wave-length in

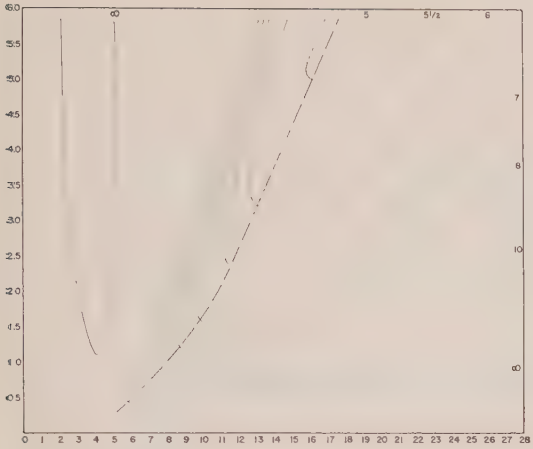


Figure 8. Stability diagram for the model presented in section 5. Abscissa is wave-length in thousands of kilometers, ordinate is vertical windshear in m sec^{-1} km^{-1}. The curves give the *e*-folding time in units of days. The dashed curve is the curve of maximum instability.

thousands of kilometers and the ordinate the mean vertical wind-shear in m sec^{-1} km^{-1}. The curve gives the *e*-folding time in days, where $T = \infty$ corresponds to neutral stability. The diagram was computed for the values $U_3 =$ 10 m sec^{-1}, $q_2 = 1.3 \times 10^{-12}$ m^{-2}, $q_4 = 5.3 \times 10^{-12}$ m^{-2}, $\beta = 16 \times 10^{-12}$ m^{-1} sec^{-1} and $U' = U'' = U_T$. The dashed line gives the curve of maximum instability. It is first of all noted from the diagram that a narrow region of instability occurs for wave-lengths around 3,000—4,000 km. This region is of no special interest since the present model does not apply on this small scale. Apart from this region the small scale is stable. A larger and larger wave-length is required to produce instability as the vertical wind-shear increases. The wave-length of maximum instability increases from about wave-number 4 ($L = 7{,}000$ km) for a shear of 0.5 m sec^{-1} km^{-1} to about wave-number 2 ($L = 14{,}000$ km) for a shear of 4 m sec^{-1} km^{-1}. The shortest *e*-folding time appears for large values of dU/dz. In the meteorological range of the vertical wind-shear we find about 5 days to be the shortest *e*-folding time. Several other diagrams using different values of the parameters U_3, q_2, q_4, and β were computed. They show only minor variations in the meteorological range of the parameters.

The main result obtained in this section is therefore that the addition of the term measuring the advection of relative vorticity stabilizes the shorter waves and introduces a region of maximum instability. Several diagrams of the vertical slope of the systems and the variations with pressure of the amplitude were constructed. The diagrams corresponding to wave-number 1 and 2 for $dU/dz = 3$ m sec^{-1} km^{-1} are reproduced in figures 9 and 10.

6. A general model for planetary waves

The study described in the earlier sections is naturally of a preliminary nature, and it is difficult to draw any final conclusions regarding the reality of the wave-solutions. However, the results obtained so far are sufficiently encouraging to warrant further investigations. In this section we shall describe a model atmosphere which can be used for further tests of the ideas presented in this paper. It is obvious a priori that no realistic prediction can be made of the planetary scales without an incorporation of heat sources and sinks and of topographical and frictional effects. However, since the author does not know of any realistic method by which heat sources can be incorporated they will be disregarded for the moment. A special problem arises in a forecast model like the one described in the previous paragraphs because only the thickness fields are prognosticated. This difficulty disappears, however, when the lower boundary condition is formulated in a more realistic sense. We shall discuss a general model for planetary flow with a four-layer model as an example. The generalization to a greater resolution will be obvious.

The atmosphere is divided into layers by levels numbered 0 to 10, where level 0 is the top of the atmosphere and 10 is the ground. We shall suppose that the levels are isobaric. The vertical velocity will be computed at the even levels: ω_2, ω_4, ω_6, ω_8, and ω_{10}. By applying the vorticity equation at the odd levels we obtain

$$\begin{aligned}
\omega_2 &= P/f_0 \cdot \mathbf{v}_1 \cdot \nabla (\zeta_1 + f) \\
\omega_4 &= \omega_2 + P/f_0 \cdot \mathbf{v}_3 \cdot \nabla (\zeta_3 + f) \\
\omega_6 &= \omega_4 + P/f_0 \cdot \mathbf{v}_5 \cdot \nabla (\zeta_5 + f) \\
\omega_8 &= \omega_6 + P/f_0 \cdot \mathbf{v}_7 \cdot \nabla (\zeta_7 + f) \\
\omega_{10} &= \omega_8 + P/f_0 \cdot \mathbf{v}_9 \cdot \nabla (\zeta_9 + f),
\end{aligned} \tag{6.1}$$

where $P = 20$ cb. Knowing the height fields at the odd levels (using the geostrophic assumption) we can compute $\omega_2, \ldots, \omega_{10}$. We are next going to forecast the thickness of four layers. Denoting in general:

$$\varphi^{(k)} = \varphi_{k-1} - \varphi_{k+1} \tag{6.2}$$

we obtain by applying the thermodynamic equation at the even levels:

$$\begin{aligned}
&\frac{\partial \varphi^{(2)}}{\partial t} + \mathbf{v}_3 \cdot \nabla \varphi^{(2)} - P\sigma_2\omega_2 = 0 \\
&\frac{\partial \varphi^{(4)}}{\partial t} + \mathbf{v}_5 \cdot \nabla \varphi^{(4)} - P\sigma_4\omega_4 = 0 \\
&\frac{\partial \varphi^{(6)}}{\partial t} + \mathbf{v}_7 \cdot \nabla \varphi^{(6)} - P\sigma_6\omega_6 = 0 \\
&\frac{\partial \varphi^{(8)}}{\partial t} + \mathbf{v}_9 \cdot \nabla \varphi^{(8)} - P\sigma_8\omega_8 = 0.
\end{aligned} \tag{6.3}$$

It is tentatively assumed than the stability factors, σ, may be assumed constant for each level. If serious errors are introduced by this procedure, we can avoid the assumption and develop a special prognostic equation for σ. This prognostic equation is obtained from the definition of σ, i.e.,

$$\sigma = -\alpha \frac{\partial \ln \Theta}{\partial p} = \frac{\partial^2 \varphi}{\partial p^2} + \frac{C_v}{C_p} \frac{1}{p} \frac{\partial \varphi}{\partial p} \tag{6.4}$$

and the thermodynamic equation. The equation is

$$\frac{\partial \sigma}{\partial t} + \mathbf{v} \cdot \nabla \sigma + \frac{\partial \sigma \omega}{\partial p} + \frac{C_v}{C_p} \frac{\omega \sigma}{p} = 0. \tag{6.5}$$

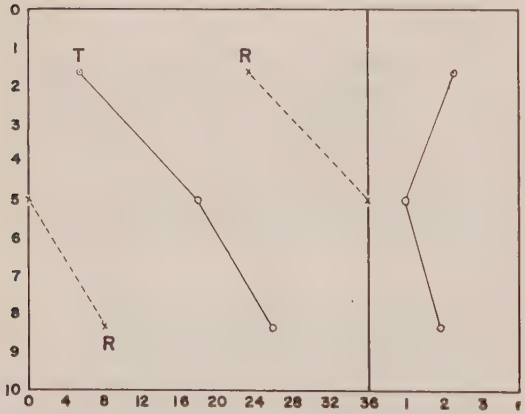


Figure 9. The vertical slope of troughs and ridges and the vertical variation of amplitude for the zonal wave-number, $N = 1$, computed from the model presented in section 5.

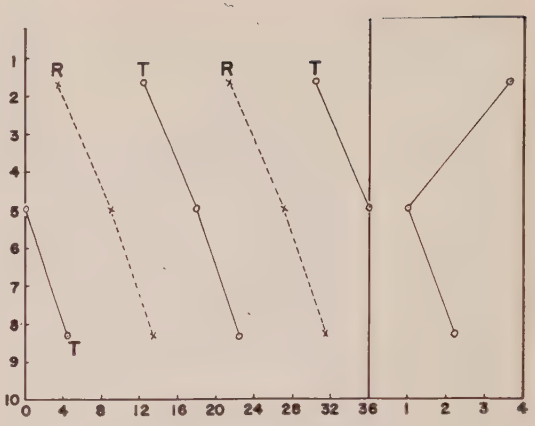


Figure 10. As figure 9 with $N = 2$.

When the prognostic equations (6.3) have been used to extrapolate the thickness fields we need to reconstruct the windfield at the odd levels. This can be done from the geostrophic assumption if we know the windfield at one level. The desired information can be obtained from the lower boundary condition where we have approximately

$$\begin{aligned}
\omega_{10} = \left(\frac{dp}{dt}\right)_{10} \approx \left(\frac{\partial p}{\partial t}\right)_{10} + \\
+ w_{10}\left(\frac{\partial p}{\partial z}\right)_{10} \approx \varrho_{10}\frac{\partial \varphi_{10}}{\partial t} - g\varrho_{10}\,\omega_{10}
\end{aligned} \tag{6.6}$$

or

$$\frac{\partial \varphi_{10}}{\partial t} = \frac{\omega_{10}}{\varrho_{10}} + g\,w_{10} \tag{6.7}$$

ω_{10} is known from (6.1) and w_{10} can be approximated as described by the author (1961, c) or by CRESSMAN (1960) from a consideration of friction and mountains.

The windfield, u_{10} and v_{10}, at the new time-step may be computed from φ_{10}. Interpolating we obtain

$$\begin{aligned}
u_9 &= u_{10} + \frac{1}{2}u^{(8)} \\
v_9 &= v_{10} + \frac{1}{2}v^{(8)}
\end{aligned} \tag{6.8}$$

The wind-fields at the odd level can now be computed from $\mathbf{v}_9$ and the thermal winds $\mathbf{v}^{(8)}$, $\mathbf{v}^{(6)}$, $\mathbf{v}^{(4)}$, and $\mathbf{v}^{(2)}$, and we are ready for a new time-step which will start with a computation of the vertical velocities.

A few comments will be made about the model outlined above. The model has been formulated with a use of the geostrophic assumption for advection purposes. Due to inconsistencies connected with the use of the geostrophic assumption it may be necessary to make systematic use of the balance equation, at least initially. It is quite likely that internal friction and heat sources and sinks will have to be included. Since we lack a good theoretical background, it may be necessary to start with empirical expressions.

It is realized that the special use of the lower boundary conditions made in the model introduces external gravity waves into the model, contrary to the theoretical model treated in the earlier sections.

A special problem arises in connection with the initial data. From the assumptions made from the outset it follows that only planetary scales shall be present in the initial data. One may isolate the planetary flow in a number of ways. The most obvious is to make an analysis of the initial fields in spherical harmonic functions and use only the first harmonics. If this method is selected it may turn out to an advantage to perform the whole integration using a spectral form of the equations.

7. Summary and discussions

The present study of the internal dynamics of the planetary wave motion is an attempt to investigate the structure and stability of these waves in a simple three-parameter model constructed in accordance with the results of the studies by BURGER (1958) and WELANDER (1960). These studies concluded that the geostrophic assumption is most valid on the planetary scale, and that the vorticity equation assumes a stationary character for this scale. We have found that the application of these principles to a three-parameter model result in unstable planetary waves, that these waves slope toward the west with decreasing pressure, and that they transport heat toward the north.

It has been suggested (SALTZMAN, 1959) that the planetary waves receive part of their kinetic energy from the small scale motion. Since this non-linear interaction will not be present in the suggested model, we may use the model to test the validity of this suggestion.

The present study can be generalized in a number of ways. It seems at the moment important to try to include the meridional scale, heat sources and friction and also to make some preliminary non-linear integrations.

REFERENCES

BURGER, A. P., 1958: Scale Consideration of Planetary Motions of the Atmosphere, *Tellus*, Vol. **10**, pp. 195—205.

CHARNEY, J. G., 1947: The Dynamics of Long Waves in a Baroclinic Westerly Current, *Journal of Meteorology*, Vol. **4**, pp. 135—162.

CHARNEY, J. G., and ELIASSEN, A., 1949: A Numerical Method for Predicting the Perturbations of the Middle Latitude Westerlies, *Tellus*, Vol. **1**, pp. 38—55.

CRESSMAN, G. P., 1958: Barotropic Divergence and Very Long Waves, *Monthly Weather Review*, Vol. **86**, pp. 293—297.

— 1960: Improved Terrain Effects in Barotropic Forecasts, *Monthly Weather Review*, Vol. **88**, pp. 327—342.

ELIASEN, E., 1958: A Study of Long Atmospheric Waves on the Basis of Zonal Harmonic Analysis, *Tellus*, Vol. **10**, pp 206—215.

LORENZ, E. N., 1960: Maximum Simplification of the Dynamic Equations, *Tellus*, Vol. **12**, pp. 243—254.

MARTIN, D. E., 1958: An Investigation of the Systematic Errors in the Barotropic Forecasts, *Tellus*, Vol. **10**, pp. 451—465.

PEIXOTO, JOSE P., 1960: *Hemispheric Temperature Conditions During the Year 1950*, Scientific Report No. 4, Planetary Circulations Project, Mass. Inst. of Technology, 211 pp.

SALTZMAN, B., 1959: On the Maintenance of the Large-Scale Quasi-Permanent Disturbances in the Atmosphere, *Tellus*, Vol. **11**, pp. 423—431.

SMAGORINSKY, J., 1953: The Dynamical Influence of Large-Scale Heat Sources and Sinks on the Quasi-Stationary Mean Motions of the Atmosphere, *Quarterly Journal of the Royal Meteorological Society*, Vol. **97**, pp. 342—366.

WELANDER, P. R., 1960: *Theory of the Long Waves in a Zonal Atmosphere Flow*, Woods Hole Oceanographic Institution, Ref. No. 60—49, 35 pp.

WIIN-NIELSEN, A., 1959: On Barotropic and Baroclinic Models with Special Emphasis on Ultra-Long Waves, *Monthly Weather Review*, Vol. **87**, pp. 171—183.

— 1961 a: Diagnosis of Divergence in a Three-Parameter Numerical Prediction Model, *Monthly Weather Review*, Vol. **89**, pp. 67—73.

— 1961 b: A Note on the Behavior of Very Long Waves in Simple Baroclinic Models, *Journal of Meteorology*, Vol. **18**, pp. 204—208.

— 1961 c: On the Distributions of Temperature Relative to Height in Stationary Planetary Waves, *Tellus*, Vol. **13**, pp. 127—139.

A Dynamical Method of Analyzing Meteorological Data[1]

By PHILIP DUNCAN THOMPSON, Colonel, U.S. Air Force, Institute of Meteorology[2], University of Stockholm

(Manuscript received April 17, 1961)

Abstract

This paper deals with the problem of reconstructing meteorological conditions in "holes"—i.e., large regions from which no data are received, but which are surrounded by regions in which the frequency and density of observations are fairly high. The method proposed here is a combination of numerical procedures of analysis and prediction, by which one can generate a series of increasingly exact analyses in a hole. It is shown that the error of this method approaches a definite lower limit which depends on:

1) The size and shape of a hole
2) The average time required for individual particles to travel across the hole.
3) The average absolute rate of flow across the boundary of the hole, and
4) A statistical measure of the inaccuracies inherent in the prediction system.

Numerical estimates of the accuracy attainable by this method indicate that dynamical methods are significantly better than conventional methods based on pure interpolation and smoothing. Finally, some aspects of the problem of network design are reviewed in the light of our present results.

1. Introduction and General Remarks

With the advent of numerical methods of weather prediction, increasing attention and effort have been given to the development of numerical methods of weather analysis – i.e., methods by which meteorological data can be processed automatically to reconstruct the "true" state of the atmosphere at particular instants in time, generally at the international observation times. This trend originated in the need for a data-processing system whose speed is compatible with that of the prediction system or, ultimately, with that of electronic computing devices. More recently, however, it has received equally strong impetus from the realization that numerical methods of analysis could make more complete and systematic use of an already large and increasingly unmanageable volume of data.

The earliest methods of numerical weather analysis, notably those of GILCHRIST and CRESSMAN (1954) and BERGTHORSSON and DÖÖS (1955), are essentially objective procedures of interpolation, whereby the values of a variable at a regular network of points can be computed from its measured values at irregularly spaced observation stations *and at a single time*. These methods have the additional function of "smoothing out" or reducing nonsystematic errors of observation and random roundoff errors, particularly in regions where observation

[1] Published with permission of the Commander, Air Weather Service, U.S. Air Force.

[2] Present affiliation: National Center for Atmospheric Research, Boulder, Col.

stations are fairly dense. For purposes of predicting conditions in midtroposphere, such methods of combined interpolation and smoothing have been found quite as satisfactory as conventional (and more subjective) methods of weather analysis.

Despite the success of methods of analysis based on pure interpolation and smoothing, it is apparent that they do not take full advantage of the data that are actually available. Specifically, it must be emphasized that an analysis produced by such methods is constructed from observations *made at a single time*. The human analyst, on the other hand, makes indirect use of observations at earlier times, in that his reconstruction of the atmosphere's current state is usually conditioned by his expectations or predictions based on earlier data. This is most true of skilled analysis in regions where observations are sparse, but is true to some extent even when observations are fairly dense.

The predictive aspect of the human analyst's art can be simulated numerically in a very straightforward way. To make the argument simple, let us characterize the international observing system as a number of isolated "holes" – i.e., fairly large regions from which no data are received – surrounded by regions in which observation stations are quite dense. Let us call the latter "O"-regions. We first suppose that the analysis over the O-regions has been constructed from data at some particular observation time by the usual methods of interpolation and smoothing. In the holes, we simply form a reasonable first guess, based on considerations of continuity, climatology, or the improbability of extremes. Now, beginning with the data in the O-regions and our guess at conditions in the holes, we compute a numerical forecast over *all* regions *for the next observation time*. When the next observation time comes, our prediction is simply modified to conform to the new data from the O-regions. The prediction in the holes is left unchanged, except to insure continuity across the edges of the holes. At that stage, our information consists of an analysis based on genuine data from the O-regions and a new (and presumably better) approximation to conditions in the holes. The prediction is then advanced to the next observation time, adjusted to fit the new data, and so on, to produce successively better "analyses" in the holes. In broad outlines, this "dynamical" method of analysis is substantially the same as the procedures devised by the Joint Numerical Weather Prediction Unit in Washington, the British Meteorological Office, and the International Meteorological Institute in Stockholm.

Although it is intuitively evident that dynamical methods of analysis are intrinsically superior to methods of pure interpolation and smoothing, one would still like to have some quantitative estimate of the accuracy attainable by such methods, together with some knowledge of the manner in which the accuracy depends on various controllable characteristics of the system. It is clear, for example, that the accuracy of analysis attainable in the holes depends on their size and on the accuracy inherent in the prediction system itself. But in what way are all these factors related functionally? It is possible, of course, to establish an empirical relationship by a long series of numerical experiments, in which holes of various sizes are simulated by deliberately denying access to the data from regions where, in fact, the coverage is quite good. It would still be desirable, however, to find some general and explicit mathematical relationship between the accuracy of analysis in the holes and the variables that determine it. An analytical treatment, even for hypothetical conditions, would serve as a guide in the design and interpretation of numerical experiments, and might also suggest the existence of variables whose relevance was previously unsuspected.

The purpose of this paper, to state it briefly, is simply to describe a dynamical method of analysis and to relate its accuracy to the shape and dimensions of a hole, the accuracy inherent in the scheme of prediction, and the actual regime of flow in the neighborhood of the hole. The main results indicate that the overall error in a hole tends toward a definite lower limit, which corresponds to a balance between two opposing effects. These are:

1. The effect of continually injecting "new" information into the hole. This effect always tends to *decrease* the mean-square wind error. It is proportional to the error itself, proportional to the average absolute speed of flow across the edges of the hole, and inversely

proportional to the linear dimensions of the hole.

2. The effect of inaccuracy in the basic system of prediction. This always tends to *increase* the mean-square wind error. It is proportional to the average time required for individual elements of fluid to traverse the hole, proportional to the square of the linear dimensions of the hole, and proportional to the mean-square of the individual rate of change of absolute vorticity (which, in the case under study, is the sole measure of indeterminacy).

The relationship described above provides a means of estimating the general level of accuracy attainable by a dynamical method of analysis which, in all essential respects, is identical to some of the dynamical methods of prediction now in routine use. Estimates of the mean-square wind error, based on plausible values of the various statistical factors, indicate that dynamical methods of analysis are significantly more accurate than pure interpolation and smoothing, particularly when the holes are more than a few hundred kilometers in diameter.

These results, while of obvious interest to the practitioner of numerical weather prediction, are of greatest importance in the design of networks of meteorological observing stations. Other factors being fixed, the overall error in a hole is a function of the hole's linear dimensions. Accordingly, if one specifies the margin of error that can be tolerated, it is possible to estimate how small the holes must be, in order to guarantee that the required standard of accuracy is maintained. In general, the maximum permissible hole-size is greater for dynamical methods of analysis than it is for optimum methods of interpolation and smoothing. Thus, since the cost of data-processing is a very small fraction of the cost of maintaining and operating a network of observing stations, dynamical methods of analysis are of considerable economic importance.

2. A Dynamical Method of Analysis Based on the Principle of Vorticity Conservation

The method of analysis that will be discussed in subsequent sections makes use of a numerical method of prediction that has been in routine use since 1955. This method is based on the so-called principle of conservation of absolute vorticity, which simply states that the absolute vorticity of each individual fluid element is the same at all points along the horizontal projection of its trajectory. Strictly speaking, this principle applies only to flows which are nondivergent in near-horizontal surfaces; however, the real motion at levels in midtroposphere is sufficiently nondivergent that vorticity is approximately conserved at those levels. Numerical methods of prediction based on this principle have so far proved to be about as accurate as more general and complicated methods, and will probably serve as a standard for some time to come.

The principles underlying this method of analysis are formulated mathematically in the following equations:

$$\frac{d\eta}{dt}=0 \tag{1}$$

$$\nabla\cdot\mathbf{v}=0 \tag{2}$$

in which $d/dt=\partial/\partial t+\mathbf{v}\cdot\nabla$; t is elapsed time; $\mathbf{v}$ is the horizontal projection of the velocity vector; ∇ is the horizontal vector gradient; η, the absolute vorticity, is $f+\mathbf{k}\cdot\nabla\times\mathbf{v}$; f is the vertical component of the earth's absolute vorticity; $\mathbf{k}\cdot\nabla\times\mathbf{v}$ and $\nabla\cdot\mathbf{v}$ are the relative vorticity and horizontal divergence of $\mathbf{v}$, respectively; and $\mathbf{k}$ is a unit vector directed vertically upward. From Eq. (2), we infer that there exists a streamfunction ψ, such that

$$\mathbf{v}=\mathbf{k}\times\nabla\psi \tag{3}$$

from which it follows that

$$\begin{aligned}\mathbf{k}\cdot\nabla\times\mathbf{v}&=\nabla^2\psi\\ \eta&=f+\nabla^2\psi\end{aligned} \tag{4}$$

Eqs. (1), (3) and (4) provide a means of predicting ψ at all points and times if ψ is known *initially* at all points.

Now, let us suppose that ψ is periodically measured at all points in a large region O, surrounding an isolated region A from which no data are received (i.e., an isolated hole). See Fig. 1. We also imagine that, at some arbitrarily chosen observation time $t=0$, we have formed a guess at the distribution of ψ in A. If A is not very large, it is reasonable to guess that the vorticity in A is uniform and equal to its average over A. Thus, a good first guess ψ_0 is defined by

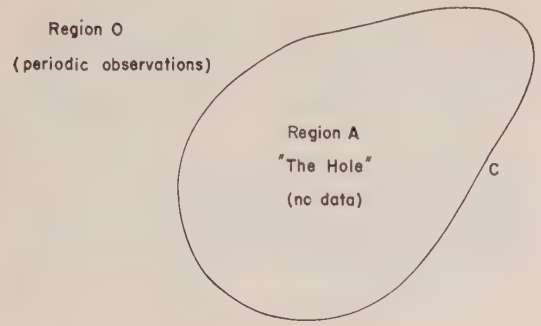


Fig. 1. Schematic representation of a "hole".

$$\nabla^2\psi_0 = \frac{1}{A}\int_A \nabla^2\psi \, dA = \frac{1}{A}\oint_C \frac{\partial\psi}{\partial n}\, ds \qquad (5)$$

where A is the area of the region A, dA is an area increment of A, a derivative with respect to n is directed normal (outward) to the boundary C of the region A, and ds is a length increment of C. Since $\partial\psi/\partial n$ can be computed along C (by hypothesis), it is possible to calculate the average vorticity over A. Moreover, since ψ is presumed known on C, the distribution of ψ_0 can be computed by solving Eq. (5) for known values of ψ_0 on C. This equation is of the Poisson type, and there are a variety of well-known numerical methods for solving it.

To summarize the situation at the arbitrarily chosen observation time $t=0$, we possess measurements of ψ at all points in O and have approximate values of ψ at all points in the hole A. We next compute the values of η at $t=0$ in both O and A from Eq. (4) and compute the velocity $\mathbf{v}$ at $t=0$ from Eq. (3). Using the latter, we compute the trajectories of all fluid elements in O and A over a very short period of time Δt. According to Eq. (1), the value of η at the end of each trajectory is the same as the value of η at the beginning. Thus, the values of η at $t=\Delta t$ can be predicted at all points in both O and A. Moreover, $\nabla^2\psi$ at $t=\Delta t$ is given by Eq. (4), and ψ at $t=\Delta t$ can be computed from $\nabla^2\psi$ at $t=\Delta t$. This procedure generates a new distribution of ψ in both O and A at $t=\Delta t$ (not, to be sure, the correct distribution), and can therefore be repeated again and again until the prediction reaches the next observation time.

At the next observation time $t=N\Delta t$, we receive new measurements of ψ in O, and the predicted values of ψ in O are replaced by the correct values. In the region A, we retain the predicted (and still incorrect) distribution of $\nabla^2\psi$, but discard the *predicted* distribution of ψ. For the latter we substitute a distribution of ψ that is computed from the predicted distribution of $\nabla^2\psi$ in A and from the newly observed values of ψ on C.

At $t=N\Delta t$ we not only possess genuine observations of ψ in O, but have predicted an approximate distribution of vorticity in A that depends partly on our initial guess in A and partly on an initially correct distribution of vorticity to the windward side of A. The presumption, therefore, is that the distribution of $\nabla^2\psi$ in A (or ψ in A) is better at $t=N\Delta t$ than it was at $t=0$. If this is indeed true, we have only to repeat the process over the interval from $t=N\Delta t$ to $t=2\,N\Delta t$, and so on, in order to reduce the error in A to some minimum value. The real question, of course, is whether or not the overall error of ψ in the hole does approach a definite minimum value as the process is carried on over a long period of time. If so, moreover, what determines the limiting error of ψ in the hole?

3. Formulation and Analysis of the Mathematical Problem

The problems just posed will be dealt with in only one extreme case, namely when the frequency and density of observations in the region surrounding an isolated hole are great enough that we may regard the true state outside the hole as known at all points and at all times. This restriction does not permit us to study one very interesting and relevant question, concerning the way in which the overall accuracy in the hole depends on the frequency of observation outside the hole; it is, however, possible to attain high frequency of observation in practice, simply because the major cost of upper-air observations is not the cost of expended balloons and unretrieved radiosonde instruments, but rather the cost of maintaining staff and facilities, plus administrative overhead.

From the standpoint of the mathematical analysis, the important consequence of the restriction stated above is that the error of

analysis can be assumed to be zero on the boundary of a hole, as well as at all points outside the hole. The error of analysis inside the hole will be different from zero, of course, because the initial guess in the hole is generally in error.

We first define a statistical measure E of the analysis error inside the hole:

$$E = \frac{1}{A}\int_A \nabla\varepsilon \cdot \nabla\varepsilon \, dA \tag{6}$$

in which ε is the analysis error in the hole. That is,

$$\varepsilon = \psi_a - \psi$$

where ψ is the true streamfunction and ψ_a is the approximate streamfunction computed by the dynamical method outlined in Section 2. The remaining notation in Eq. (6) is that defined in Section 2. Thus, since the flow is approximately nondivergent, E is essentially the mean-square vector wind error, averaged over the hole. Our main concern is to find how time-variations of E depend on hole-size, the statistical properties of flow in the hole and other statistical factors.

Differentiating E with respect to time, we find that

$$\frac{\partial E}{\partial t} = \frac{2}{A}\int_A \nabla\varepsilon \cdot \nabla\frac{\partial\varepsilon}{\partial t}\, dA =$$

$$\frac{2}{A}\int_A \left(\nabla\cdot\varepsilon\nabla\frac{\partial\varepsilon}{\partial t} - \varepsilon\nabla^2\frac{\partial\varepsilon}{\partial t}\right) dA$$

or, by making use of Gauss' theorem,

$$\frac{\partial E}{\partial t} = \frac{2}{A}\oint_C \varepsilon\frac{\partial}{\partial n}\left(\frac{\partial\varepsilon}{\partial t}\right) ds - \frac{2}{A}\int_A \varepsilon\nabla^2\frac{\partial\varepsilon}{\partial t}\, dA$$

where the notation is that defined in Section 2. The first term on the right-hand side of the equation above vanishes, because ε vanishes at all points on C. Thus

$$\frac{\partial E}{\partial t} = -\frac{2}{A}\int_A \varepsilon\nabla^2\frac{\partial\varepsilon}{\partial t}\, dA \tag{7}$$

In order to calculate $\partial E/\partial t$, therefore, it remains to find how the analysis error ε varies with time.

The equation governing the analysis error is derived from the equation that describes the procedure for computing the approximate streamfunction ψ_a – i.e., the equation for conservation of absolute vorticity. Since f does not depend on time, Eq. (1) may be written as:

$$\frac{\partial}{\partial t}\nabla^2\psi_a + \mathbf{k}\cdot\nabla\psi_a \times \nabla(f + \nabla^2\psi_a) = 0$$

The true streamfunction ψ, on the other hand, is governed by the equation

$$\frac{\partial}{\partial t}\nabla^2\psi + \mathbf{k}\cdot\nabla\psi\times\nabla(f+\nabla^2\psi) = \frac{d\eta}{\partial t} \equiv S$$

Subtracting the equation above from the one preceding it we now see that

$$\frac{\partial}{\partial t}\nabla^2\varepsilon + \mathbf{k}\cdot\nabla\psi_a\times\nabla(f+\nabla^2\psi_a) - \mathbf{k}\cdot\nabla\psi\times$$
$$\times\nabla(f+\nabla^2\psi) = -S$$

Moreover, adding and subtracting $\mathbf{k}\cdot\nabla\psi\times$ $\times\nabla(f+\nabla^2\psi_a)$ and factoring, we may rewrite this equation as

$$\frac{\partial}{\partial t}\nabla^2\varepsilon + \mathbf{k}\cdot\nabla\psi\times\nabla(\nabla^2\psi_a - \nabla^2\psi) + \mathbf{k}\cdot$$
$$\cdot\nabla(\psi_a - \psi)\times\nabla(f+\nabla^2\psi_a) = -S$$

or, since $\mathbf{k}\cdot\nabla\psi\times\nabla = \mathbf{v}\cdot\nabla$

$$\frac{d}{dt}\nabla^2\varepsilon + \mathbf{k}\cdot\nabla\varepsilon\times\nabla(f+\nabla^2\psi_a) = -S \tag{8}$$

The equation above is now multiplied by ε and integrated over A, with the result that

$$\int_A \varepsilon\frac{d}{dt}\nabla^2\varepsilon\, dA + \int_A \mathbf{k}\cdot\nabla\left(\frac{\varepsilon^2}{2}\right)\times$$
$$\nabla(f+\nabla^2\psi_a)\, dA = -\int_A \varepsilon S\, dA$$

The second term on the left hand side of this equation may be transformed with the aid of Stokes' theorem, as follows:

$$\int_A \mathbf{k}\cdot\nabla\left(\frac{\varepsilon^2}{2}\right)\times\nabla(f+\nabla^2\psi_a)\, dA = \int_A \mathbf{k}\cdot\nabla\times$$
$$\times\left[\frac{\varepsilon^2}{2}\nabla(f+\nabla^2\psi_a)\right] dA = \oint_C \frac{\varepsilon^2}{2}\frac{\partial}{\partial s}(f+\nabla^2\psi_a)\, ds$$

where a derivative with respect to s is directed along C. Since ε vanishes at all points on C, this term vanishes. Thus, the preceding equation reduces to

$$\int_A \varepsilon \frac{d}{dt} \nabla^2 \varepsilon \, dA + \int_A \varepsilon \, S \, dA = 0 \qquad (9)$$

This equation governs the analysis error ε in the hole A.

The basic equation for the changes of the overall error E is obtained by multiplying Eq. (9) by $2/A$, and by adding it to Eq. (7). Thus,

$$\frac{\partial E}{\partial t} = \frac{2}{A} \int_A \varepsilon \left(\frac{d}{dt} \nabla^2 \varepsilon - \frac{\partial}{\partial t} \nabla^2 \varepsilon \right) dA + \frac{2}{A} \int_A \varepsilon \, S \, dA \qquad (10)$$

Now, it should be noted that the quantity in parentheses is just the "horizontal advection" of $\nabla^2 \varepsilon$. According to the method outlined in Section 2, the horizontal advection of vorticity is to be carried out by conserving vorticity along the particle trajectories[1], i.e.,

$$\frac{d}{dt} \nabla^2 \varepsilon - \frac{\partial}{\partial t} \nabla^2 \varepsilon = \lim_{\Delta t \to 0} \left[\frac{\nabla^2 \varepsilon (\mathbf{r}) - \nabla^2 \varepsilon (\mathbf{r} - \mathbf{v} \Delta t)}{\Delta t} \right]$$

where $\mathbf{r}$ is a variable position vector, corresponding to a point (x, y). Substituting this expression into Eq. (10) and reversing the order of the limiting processes, we find that

$$\frac{\partial E}{\partial t} = \lim_{\Delta t \to 0} \frac{2}{A \Delta t} \int_A \left[\varepsilon (\mathbf{r}) \nabla^2 \varepsilon (\mathbf{r}) - \varepsilon (\mathbf{r}) \nabla^2 \varepsilon (\mathbf{r} - \mathbf{v} \Delta t) \right] dA + \frac{2}{A} \int_A \varepsilon \, S \, dA \qquad (11)$$

The remaining problem is to express the right hand side of this equation in terms of statistics that are known or can be predicted.

Let us begin by considering the first term on the right hand side of Eq. (11). Denoting that quantity as I, and adding zero to it, we have

$$I = \lim_{\Delta t \to 0} \frac{2}{A \Delta t} \int_A \varepsilon (\mathbf{r}) \left[\nabla^2 \varepsilon (\mathbf{r}) - \nabla^2 \varepsilon (\mathbf{r} - \bar{\mathbf{v}} \Delta t \right] \cdot \cdot \, dA + \frac{2}{A} J \qquad (12)$$

where

$$J = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_A \varepsilon (\mathbf{r}) \left[\nabla^2 \varepsilon (\mathbf{r} - \bar{\mathbf{v}} \Delta t) - \nabla^2 \varepsilon (\mathbf{r} - \mathbf{v} \Delta t) \right] dA$$

and $\mathbf{v}$ is a mean velocity, averaged over the boundary C. Taylor-expanding $\nabla^2 \varepsilon$ in the neighborhood of $\mathbf{r}$, we may rewrite J as

$$J = \int_A \varepsilon (\mathbf{v} - \bar{\mathbf{v}}) \cdot \nabla (\nabla^2 \varepsilon) \, dA$$

or, since both $\mathbf{v}$ and $\bar{\mathbf{v}}$ are nondivergent,

$$J = \int_A \varepsilon \nabla \cdot \left[\nabla^2 \varepsilon (\mathbf{v} - \bar{\mathbf{v}}) \right] dA$$

$$= \int_A \nabla \cdot \left[\varepsilon \nabla^2 \varepsilon (\mathbf{v} - \bar{\mathbf{v}}) \right] dA - \int_A \nabla^2 \varepsilon (\mathbf{v} - \bar{\mathbf{v}}) \cdot \nabla \varepsilon \, dA$$

By making use of Gauss' theorem, the first term on the right hand side of this equation can be transformed into a line integral around C, whose integrand contains ε as a factor. Thus, since ε vanishes at all points on C, that integral also vanishes, and the preceding equation reduces to

$$J = - \int_A \nabla^2 \varepsilon (\mathbf{v} - \bar{\mathbf{v}}) \cdot \nabla \varepsilon \, dA$$

$$= - \int_A \nabla \cdot \left[(\mathbf{v} - \bar{\mathbf{v}}) \cdot \nabla \varepsilon \right] \nabla \varepsilon \, dA + \int_A \nabla \varepsilon \cdot \nabla \left[(\mathbf{v} - \bar{\mathbf{v}}) \cdot \nabla \varepsilon \right] dA$$

Expanding the gradient of a dot-product according to the rules for vector differentiation, and noting that $\bar{\mathbf{v}}$ is constant, we then find that

$$J = - \int_A \nabla \cdot \left[(\mathbf{v} - \bar{\mathbf{v}}) \cdot \nabla \varepsilon \right] \nabla \varepsilon \, dA + \int_A \nabla \varepsilon \cdot \left[(\mathbf{v} - \bar{\mathbf{v}}) \cdot \nabla \right] \nabla \varepsilon \, dA + \int_A \nabla \varepsilon \cdot (\nabla \varepsilon \cdot \nabla) \, \mathbf{v} \, dA - \int_A \nabla \varepsilon \cdot \left[(\nabla \times \mathbf{v}) \times \nabla \varepsilon \right] dA -$$

[1] A preliminary study of the one-dimensional case indicated that this method of horizontal advection is the only one which guarantees that E converges to zero when S is zero.

$$-\int_A \nabla\varepsilon \cdot [(\nabla \times \nabla\varepsilon) \times (\mathbf{v} - \overline{\mathbf{v}})]\, dA \quad (13)$$

The fourth and fifth terms on the right hand side of this equation vanish; the fourth, because $\nabla\varepsilon$ is always normal to the vector $(\nabla \times \mathbf{v}) \times \nabla\varepsilon$; and the fifth, because $\nabla \times \nabla\varepsilon$ vanishes identically. The first term may be transformed by Gauss' theorem, to give it the form

$$-\int_A \nabla \cdot [(\mathbf{v} - \overline{\mathbf{v}}) \cdot \nabla\varepsilon]\, \nabla\varepsilon\, dA =$$

$$= -\oint_C \frac{d\varepsilon}{\partial n} (\mathbf{v} - \overline{\mathbf{v}}) \cdot \nabla\varepsilon\, ds$$

or, since $\partial\varepsilon/\partial s$ vanishes on C,

$$-\int_A \nabla \cdot [(\mathbf{v} - \overline{\mathbf{v}}) \cdot \nabla\varepsilon]\, \nabla\varepsilon\, dA =$$

$$= -\oint_C (V_n - \overline{V}_n) \left(\frac{\partial\varepsilon}{\partial n}\right)^2 ds =$$

$$= -\oint_C (V_n - \overline{V}_n)\, \nabla\varepsilon \cdot \nabla\varepsilon\, ds$$

in which V_n and $\overline{V}_n$ are the components of $\mathbf{v}$ and $\overline{\mathbf{v}}$, respectively, directed normal (outward) to C. The second term on the right hand side of Eq. (13) may be written as

$$\int_A \nabla\varepsilon \cdot [(\mathbf{v} - \overline{\mathbf{v}}) \cdot \nabla]\, \nabla\varepsilon\, dA =$$

$$= \frac{1}{2} \int_A (\mathbf{v} - \overline{\mathbf{v}}) \cdot \nabla\, (\nabla\varepsilon \cdot \nabla\varepsilon)\, dA$$

or, since both $\overline{\mathbf{v}}$ and $\mathbf{v}$ are nondivergent,

$$\int_A \nabla\varepsilon \cdot [(\mathbf{v} - \overline{\mathbf{v}}) \cdot \nabla]\, \nabla\varepsilon\, dA =$$

$$= \frac{1}{2} \int_A \nabla \cdot (\nabla\varepsilon \cdot \nabla\varepsilon)\, (\mathbf{v} - \overline{\mathbf{v}})\, dA$$

$$= \frac{1}{2} \oint_C (V_n - \overline{V}_n)\, \nabla\varepsilon \cdot \nabla\varepsilon\, ds$$

Finally, incorporating these results into Eq. (13), we see that

$$J = -\frac{1}{2} \oint_C (V_n - \overline{V}_n)\, \nabla\varepsilon \cdot \nabla\varepsilon\, ds +$$

$$+ \int_A \nabla\varepsilon \cdot (\nabla\varepsilon \cdot \nabla)\, \mathbf{v}\, dA \quad (14)$$

By expressing $\mathbf{v}$ as $\mathbf{k} \times \nabla\psi$, the second term on the right hand side of Eq. (14) may be rewritten as

$$\int_A \nabla\varepsilon \cdot (\nabla\varepsilon \cdot \nabla)\, \mathbf{v}\, dA =$$

$$= \int_A \nabla\varepsilon \cdot \left(\frac{\partial\varepsilon}{\partial x}\frac{\partial \mathbf{v}}{\partial x} + \frac{\partial\varepsilon}{\partial y}\frac{\partial \mathbf{v}}{\partial y}\right) dA$$

$$= \int_A \left[\left(\frac{\partial\varepsilon}{\partial y}\right)^2 - \left(\frac{\partial\varepsilon}{\partial x}\right)^2\right] \frac{\partial^2\psi}{\partial x \partial y}\, dA$$

$$+ \int_A \frac{\partial\varepsilon}{\partial x}\frac{\partial\varepsilon}{\partial y}\left(\frac{\partial^2\psi}{\partial x^2} - \frac{\partial^2\psi}{\partial y^2}\right) dA$$

where x and y are cartesian horizontal coordinates whose orientation is arbitrary. Moreover, by introducing a new set of coordinates (X, Y), whose axes make an angle of 45° with those of the (x, y) coordinates, we may put the equation above in the form

$$\int_A \nabla\varepsilon \cdot (\nabla\varepsilon \cdot \nabla)\, \mathbf{v}\, dA =$$

$$= \int_A \left[\left(\frac{\partial\varepsilon}{\partial y}\right)^2 - \left(\frac{\partial\varepsilon}{\partial x}\right)^2\right] \frac{\partial^2\psi}{\partial x \partial y}\, dA \quad (15)$$

$$- \int_A \left[\left(\frac{\partial\varepsilon}{\partial Y}\right)^2 - \left(\frac{\partial\varepsilon}{\partial X}\right)^2\right] \frac{\partial^2\psi}{\partial X \partial Y}\, dA$$

in which

$$X = \frac{1}{\sqrt{2}}(x + y) \qquad Y = \frac{1}{\sqrt{2}}(x - y)$$

Now, the frequency distribution of vector deviations of wind from its average over a large ensemble of cases is very nearly independent of the direction of those deviations. Thus, if the hole A has about the same dimensions in all directions, the integral

$$\int_A \left[\left(\frac{\partial\varepsilon}{\partial y}\right)^2 - \left(\frac{\partial\varepsilon}{\partial x}\right)^2\right] \frac{\partial^2\psi}{\partial x \partial y}\, dA$$

is very nearly independent of the orientation of the x, y axes *in an average taken over a large*

ensemble of cases. In the ensemble average, therefore, the two terms on the right hand side of Eq. (15) cancel. With the understanding that all further conclusions will apply only to the ensemble average, Eq. (14) then reduces to

$$J=-\frac{1}{2}\oint_C (V_n-\bar{V}_n)\,\nabla\varepsilon\cdot\nabla\varepsilon\,ds$$

and Eq. (12) takes the form

$$I=\lim_{\Delta t\to 0}\frac{2}{A\Delta t}\int_A \varepsilon(\mathbf{r})\,[\nabla^2\varepsilon(\mathbf{r})- \nabla^2\varepsilon(\mathbf{r}-\bar{\mathbf{v}}\Delta t)]\,dA-\frac{1}{A}\oint_C (V_n-\bar{V}_n)\,\nabla\varepsilon\cdot\nabla\varepsilon\,ds \tag{16}$$

As a matter of fact, it will later be shown that the second term on the right hand side of this equation is much less than the first. Accordingly, if vorticity were conserved in actuality (i.e., if $S=0$), the change in the mean-square wind error E would be due almost entirely to the advection of $\nabla^2\varepsilon$ with the mean wind $\bar{\mathbf{v}}$ on the boundary C, without regard to variations of the wind field in the interior of A.

It remains to express the first integral on the right hand side of Eq. (16) as a combination of known or predictable statistics. First, by applying Gauss' theorem and noting that ε vanishes on C, we may put Eq. (16) in the form

$$I=-\lim_{\Delta t\to 0}\frac{2}{A\Delta t}\int_A \nabla\varepsilon\cdot\nabla\varepsilon\,dA-\lim_{\Delta t\to 0}\frac{2}{A\Delta t}\int_A \varepsilon(\mathbf{r})\,\nabla^2\varepsilon(\mathbf{r}-\bar{\mathbf{v}}\Delta t)\,dA-\frac{1}{A}\oint_C (V_n-\bar{V}_n)\,\nabla\varepsilon\cdot\nabla\varepsilon\,ds \tag{17}$$

The integral K appearing in the second term on the right hand side of the equation above may be written as

$$K=\int_A \varepsilon(\mathbf{r})\,\nabla^2\varepsilon(\mathbf{r}-\bar{\mathbf{v}}\Delta t)\,dA$$

$$=\iint_A \varepsilon(x,y)\,\frac{\partial^2\varepsilon}{\partial x^2}(x-\xi,y-\eta)\,dx\,dy+\iint_A \varepsilon(x,y)\,\frac{\partial^2\varepsilon}{\partial y^2}(x-\xi,y-\eta)\,dx\,dy$$

in which $\xi=\bar{u}\Delta t$ and $\eta=\bar{v}\Delta t$ are the x and y components of $\bar{\mathbf{v}}\Delta t$, respectively, and *are independent of x and y*. It is understood that the limits of integration are given by the coordinates (x_0, y_0) of points on the boundary C of A. Since the boundary of A is fixed, the equation above can be written as

$$K=-\frac{\partial}{\partial\xi}\iint_A \varepsilon(x,y)\,\frac{\partial\varepsilon}{\partial x}(x-\xi,y-\eta)\,dx\,dy-\frac{\partial}{\partial\eta}\iint_A \varepsilon(x,y)\,\frac{\partial\varepsilon}{\partial y}(x-\xi,y-\eta)\,dx\,dy$$

Thus, letting $\sigma=x-\xi$ and $\tau=y-\eta$, we find that

$$K=-\frac{\partial}{\partial\xi}\iint_{A'} \varepsilon(\sigma+\xi,\tau+\eta)\,\frac{\partial\varepsilon}{\partial\sigma}(\sigma,\tau)\,d\sigma\,d\tau-\frac{\partial}{\partial\eta}\iint_{A'} \varepsilon(\sigma+\xi,\tau+\eta)\,\frac{\partial\varepsilon}{\partial\tau}(\sigma,\tau)\,d\sigma\,d\tau$$

We next carry out the differentiations with respect to ξ and η. In doing so, one must bear in mind that the points on the boundary of A' have coordinates $(x_0-\xi,\ y_0-\eta)$, and that the limits of integration therefore depend on ξ and η. The contribution due to variations of the limits of integration vanish, however, simply because $\varepsilon\ (x_0, y_0)$ vanishes. Thus, from the equation above,

$$K=-\iint_{A'} \frac{\partial\varepsilon}{\partial\sigma}(\sigma+\xi,\tau+\eta)\,\frac{\partial\varepsilon}{\partial\sigma}(\sigma,\tau)\,d\sigma\,d\tau-\iint_{A'} \frac{\partial\varepsilon}{\partial\tau}(\sigma+\xi,\tau+\eta)\,\frac{\partial\varepsilon}{\partial\tau}(\sigma,\tau)\,d\sigma\,d\tau$$

or, transforming back to the (x, y) coordinates,

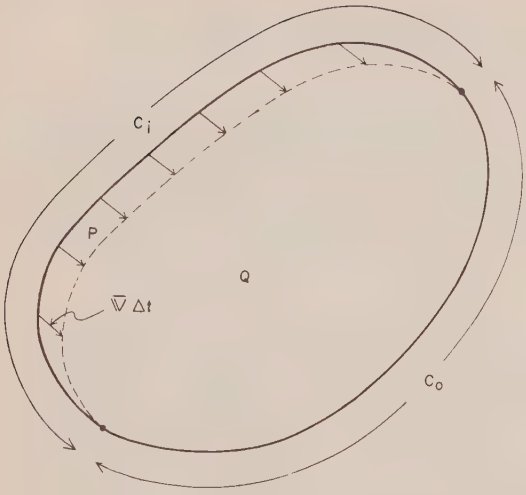


Fig. 2. Map of the region A, showing the subregions P and Q and the segments C_i and C_0 along which $\overline{\mathbf{v}}$ is directed inward and outward across C.

$$K=-\iint_A \frac{\partial\varepsilon}{\partial x}(x,y)\,\frac{\partial\varepsilon}{\partial x}(x-\xi,y-\eta)\,dx\,dy$$

$$=\iint_A \frac{\partial\varepsilon}{\partial y}(x,y)\,\frac{\partial\varepsilon}{\partial y}(x-\xi,y-\eta)\,dx\,dy$$

$$=-\int_A \nabla\varepsilon(\mathbf{r})\cdot\nabla\varepsilon(\mathbf{r}-\overline{\mathbf{v}}\Delta t)\,dA$$

Let us now distinguish between two subregions of A. One subregion P is such that all points $\mathbf{r}$ in P correspond to points $(\mathbf{r}-\overline{\mathbf{v}}\Delta t)$ *outside* A, whereas Q is such that all points $\mathbf{r}$ in Q correspond to points $(\mathbf{r}-\overline{\mathbf{v}}\Delta t)$ *inside* A, as shown in Fig. 2. Accordingly, since $\nabla\varepsilon\cdot(\mathbf{r}-\overline{\mathbf{v}}\Delta t)$ vanishes for all points $\mathbf{r}$ in P, the equation above reduces to

$$K=-\int_Q \nabla\varepsilon(\mathbf{r})\cdot\nabla\varepsilon(\mathbf{r}-\overline{\mathbf{v}}\Delta t)\,dA$$

Finally, substituting this result into the second term on the right hand side of Eq. (17), we find that

$$I=-\lim_{\Delta t\to 0}\frac{2}{A\Delta t}\int_P \nabla\varepsilon\cdot\nabla\varepsilon\,dA$$

$$-\lim_{\Delta t\to 0}\frac{2}{A\Delta t}\int_Q \nabla\varepsilon(\mathbf{r})\cdot$$

$$\cdot\left[\nabla\varepsilon(\mathbf{r})-\nabla\varepsilon(\mathbf{r}-\overline{\mathbf{v}}\Delta t)\right]dA$$

$$-\frac{1}{A}\oint_C (V_n-\overline{V}_n)\,\nabla\varepsilon\cdot\nabla\varepsilon\,ds \qquad (18)$$

The first term on the right hand side of this equation may be written as

$$-\lim_{\Delta t\to 0}\frac{2}{\overline{A}}\cdot\frac{P}{\Delta t}\cdot\frac{1}{P}\int_P \nabla\varepsilon\cdot\nabla\varepsilon\,dA \qquad (19)$$

$$=-\frac{2}{A}\left(\int_{C_i}\overline{V}_i\,ds\right)\cdot\left(\frac{1}{C_i}\int_{C_i}\nabla\varepsilon\cdot\nabla\varepsilon\,ds\right)$$

where P also denotes the area of P, C_i denotes that segment of C across which $\overline{\mathbf{v}}$ is directed inward, and $\overline{V}_i$ is the component of $\overline{\mathbf{v}}$ normal to C_i. The second term of Eq. (18) is, by definition,

$$-\lim_{\Delta t\to 0}\frac{2}{A\Delta t}\int_Q \frac{\partial\varepsilon}{\partial x}(\mathbf{r})\left[\frac{\partial\varepsilon}{\partial x}(\mathbf{r})-\right.$$

$$\left.-\frac{\partial\varepsilon}{\partial x}(\mathbf{r}-\overline{\mathbf{v}}\Delta t)\right]dA$$

$$-\lim_{\Delta t\to 0}\frac{2}{A\Delta t}\int_Q \frac{\partial\varepsilon}{\partial y}(\mathbf{r})\left[\frac{\partial\varepsilon}{\partial y}(\mathbf{r})-\right.$$

$$\left.-\frac{\partial\varepsilon}{\partial y}(\mathbf{r}-\overline{\mathbf{v}}\Delta t)\right]dA$$

or, since $\partial\varepsilon/\partial x$ and $\partial\varepsilon/\partial y$ are continuous in A, that term also equals

$$-\frac{2}{A}\int_A\left(\frac{\partial\varepsilon}{\partial x}\overline{\mathbf{v}}\cdot\nabla\frac{\partial\varepsilon}{\partial x}+\frac{\partial\varepsilon}{\partial y}\overline{\mathbf{v}}\cdot\nabla\frac{\partial\varepsilon}{\partial y}\right)dA$$

$$=-\frac{2}{A}\int_A\overline{\mathbf{v}}\cdot\nabla\left(\frac{\nabla\varepsilon\cdot\nabla\varepsilon}{2}\right)dA$$

$$=-\frac{1}{A}\int_A\nabla\cdot(\nabla\varepsilon\cdot\nabla\varepsilon)\,\overline{\mathbf{v}}\,dA$$

$$=-\frac{1}{A}\oint_C\overline{V}_n\,\nabla\varepsilon\cdot\nabla\varepsilon\,ds$$

Substituting this expression for the second term of Eq. (18) and substituting from Eq. (19) into the first term, we see that

$$I=-\frac{2}{A}\left(\oint_{C_i}\overline{V}_i\,ds\right)\left(\frac{1}{C_i}\oint_{C_i}\nabla\varepsilon\cdot\nabla\varepsilon\,ds\right)$$

$$-\frac{1}{A}\oint_C \overline{V}_n \nabla\varepsilon\cdot\nabla\varepsilon\, ds - \frac{1}{A}\oint_C (V_n-\overline{V}_n)\nabla\varepsilon\cdot\nabla\varepsilon\, ds \tag{20}$$

We can now verify that variations of $\mathbf{v}$ in A have little effect on the value of I. This follows from the fact that

$$\oint_C (v_n-\overline{v}_n)\, ds = 0$$

by definition. Thus, since $\nabla\varepsilon\cdot\nabla\varepsilon$ is always positive and more or less uniform along C, the third term of Eq. (20) is generally much less than the second – which, in turn, is much smaller than the first. The second term, moreover, may be written as

$$\frac{1}{A}\int_{C_i}\overline{V}_i\nabla\varepsilon\cdot\nabla\varepsilon\, ds - \frac{1}{A}\int_{C_0}\overline{V}_0\nabla\varepsilon\cdot\nabla\varepsilon\, ds$$

where C_0 denotes that segment of C across which the flow is outward, and $\overline{\mathbf{v}}_0$ is the component of $\mathbf{v}$ normal to C_0. Eq. (20) then reduces to

$$I=-\frac{2C_i}{A}\left(\frac{1}{C_i}\int_{C_i}\overline{V}_i ds\right)\left(\frac{1}{C_i}\int_{C_i}\nabla\varepsilon\cdot\nabla\varepsilon\, ds\right) + \frac{1}{A}\int_{C_i}\overline{V}_i\nabla\varepsilon\cdot\nabla\varepsilon\, ds - \frac{1}{A}\int_{C_0}\overline{V}_0\nabla\varepsilon\cdot\nabla\varepsilon\, ds$$

Noting that $\overline{v}_i$, $\overline{v}_0$ and $\nabla\varepsilon\cdot\nabla\varepsilon$ are all positive, we next approximate the second and third terms of this equation (whose sum is small), as follows:

$$\begin{aligned} I \simeq & -\frac{2C_i}{A}\left(\frac{1}{C_i}\int_{C_i}\overline{V}_i ds\right)\left(\frac{1}{C_i}\int_{C_i}\nabla\varepsilon\cdot\nabla\varepsilon\, ds\right) + \\ & + \frac{C_i}{A}\left(\frac{1}{C_i}\int_{C_i}\overline{V}_i ds\right)\left(\frac{1}{C_i}\int_{C_i}\nabla\varepsilon\cdot\nabla\varepsilon\, ds\right) - \\ & - \frac{C_0}{A}\left(\frac{1}{C_0}\int_{C_0}\overline{V}_0 ds\right)\left(\frac{1}{C_0}\int_{C_0}\nabla\varepsilon\cdot\nabla\varepsilon\, ds\right) \\ \simeq & -\frac{1}{A}\left[C_i\left(\frac{1}{C_i}\int_{C_i}\overline{V}_i ds\right)\left(\frac{1}{C_i}\int_{C_i}\nabla\varepsilon\cdot\nabla\varepsilon\, ds\right) + \right. \\ & \left. + C_0\left(\frac{1}{C_0}\int_{C_0}\overline{V}_0 ds\right)\left(\frac{1}{C_0}\int_{C_0}\nabla\varepsilon\cdot\nabla\varepsilon\, ds\right)\right] \end{aligned} \tag{21}$$

Moreover, since $\overline{\mathbf{v}}$ is nondivergent,

$$\int_A \nabla\cdot\overline{\mathbf{v}}\, dA = \oint_C \overline{V}_n\, ds = \int_{C_0}\overline{V}_0 ds - \int_{C_i}\overline{V}_i ds$$
$$= C_0\left(\frac{1}{C_0}\int_{C_0}\overline{V}_0 ds\right) - C_i\left(\frac{1}{C_i}\int_{C_i}\overline{V}_i ds\right) = 0$$

Now, in the ensemble average, C_0 is very nearly equal to C_i. From the equation above, therefore, it also follows that

$$\frac{1}{C_0}\int_{C_0}\overline{V}_0 ds \simeq \frac{1}{C_i}\int_{C_i}\overline{V}_i ds$$

Introducing these results into Eq. (21), we have

$$I=-\frac{C}{A}\left(\frac{1}{C}\oint_C|\overline{V}_n|\, ds\right)\left(\frac{1}{C}\oint_C\nabla\varepsilon\cdot\nabla\varepsilon\, ds\right) \tag{22}$$

in which C is also the circumference of A. This equation explicitly relates I to the shape and dimensions of the hole A, the average rate of flow across the boundary of the hole, and the mean-square wind error around the boundary.

On retracing our way through the foregoing derivation, we see that the term I in Eq. (11) represents the effect of advecting correct values of vorticity into A across its windward edge. Since Eq. (22) shows that I is always negative, this effect always tends to decrease the mean square wind error E. Although this result is not very surprising, it should be emphasized that it depends on the fact that the advection of vorticity is carried out by conserving vorticity along forward-extrapolated trajectories – rather than by the usual centered-difference methods. Ultimately, in fact, the success of the present method of analysis depends on this particular method of advection.

4. The Case of No Vorticity Production

If the absolute vorticity were conserved in actuality, $S=0$ and Eq. (11) reduces to

$$\frac{\partial E}{\partial t} = -\frac{CV}{A}\cdot\frac{1}{C}\oint_C\nabla\varepsilon\cdot\nabla\varepsilon\, ds \tag{23}$$

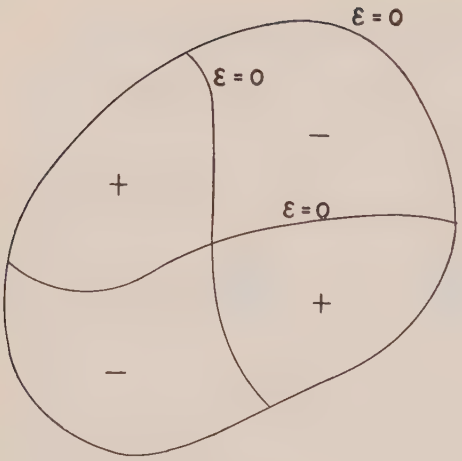


Fig. 3. A typical distribution of error ε in a hole.

in which V without subscript is defined as

$$V = \frac{1}{C} \oint_C |\bar{V}_n| \, ds$$

Now, since ε vanishes on C, the distribution of ε in A must have a cellular structure – that is, it consists of a number of cells of alternately positive and negative error separated by lines of zero error, as shown schematically in Fig. 3. It is reasonable to suppose that the mean-square wind error around C is about the same as that averaged around any single cell and, moreover, that the mean-square error over the whole of A is about the same as that averaged over a single cell. Thus, assuming that ε has a typical distribution in any single cell, we suppose that

$$\frac{1}{C} \oint_C \nabla\varepsilon \cdot \nabla\varepsilon \, ds = \lambda E$$

where λ is assumed to be a constant, depending on the characteristic distribution of ε in a single cell, and whose magnitude is about unity. With this hypothesis, Eq. (23) takes the form

$$\frac{\partial E}{\partial t} = -\frac{\lambda C V E}{A} \qquad (24)$$

C and A are the circumference and area of the hole and are independent of time. Moreover, V is independent of time in an ensemble average, whence the solution of Eq. (24) is approximately

$$E = E_0 \exp\left(-\frac{\lambda C V t}{A}\right) \qquad (25)$$

in which E_0 is the value of E at $t = 0$. This equation implies that, if vorticity were actually conserved, the mean-square wind error E would converge asymptotically to zero. Accordingly, since E vanishes if and only if ε vanishes everywhere in A, the analysis error in A would presumably converge to zero also.

This conclusion has been verified by a series of numerical experiments carried out by N. N. Richardson (1961). In these experiments, observed values of ψ were simulated by its predicted values, computed by the principle of vorticity conservation from initial data given over *all* regions. The predicted values of ψ were, in effect, the values that would have been observed if vorticity were actually conserved. Having generated a sequence of "hypothetical" observations of ψ over *all* regions, Richardson then pretended that they were not available over square "holes" of various sizes, and computed successive approximations to ψ in the holes by the dynamical method outlined in Section 2. His results in this case are summarized in Figs. 4(a) and 4(b).

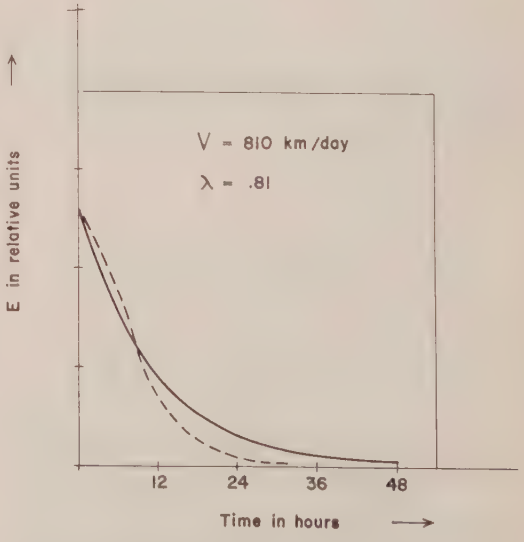


Fig. 4(a). Theoretical (solid curve) and observed (dashed curve) decrease of E for a square hole 1,200 km × 1,200 km. After Richardson.

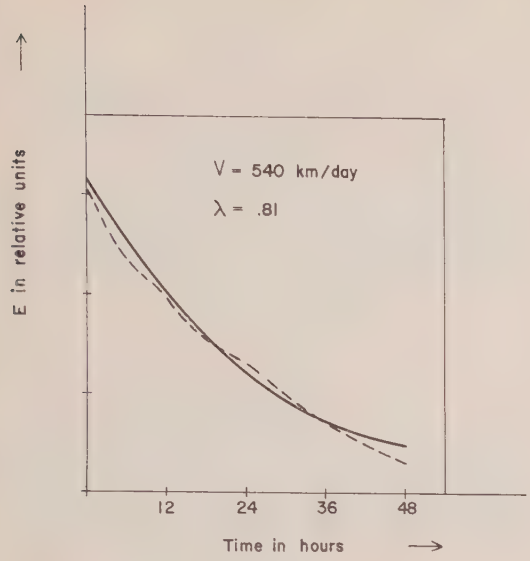


Fig. 4(b). Theoretical (solid curve) and observed (dashed curve) decrease of E for a square hole 1,800 km × 1,800 km. After Richardson.

The dashed curve on Fig. 4(a) shows how E/E_0 was found to decrease with time for a hole 1,200 km across. (It should be noted that the "correct" values of ψ in A were actually available to compute E.) The solid curve, on the other hand, shows the decrease of E/E_0 predicted from Eq. (25), for the observed initial value of V. The value of λ, which was chosen to give the best fit *in this one case*, was 0.81. Fig. 4(b) shows the same type of result for a square hole 1,800 km across. In this case, the predicted decrease of E/E_0 was computed from the observed initial value of V and the optimum value of λ for the *previous* case.

The most striking feature of Richardson's results is that the mean-square wind error apparently does converge to zero, as predicted by Eq. (25). Moreover, the manner in which E is predicted to decrease agrees very well with the result of numerical experiment. This strongly suggests, therefore, that Eq. (24) is very nearly valid in the present case, and that Eq. (11) for the more general case has the form

$$\frac{\partial E}{\partial t} = -\frac{\lambda C V E}{A} + \frac{2}{A}\int_A \varepsilon S \, dA \qquad (26)$$

The formula above, as we shall see later, provides a means of estimating the overall accuracy attainable by the general method of analysis described earlier.

5. The Limiting Error in the General Case

So far, we have analyzed only the error-diminishing effects due to advecting correct values of vorticity across the windward edge of a hole, and it remains to investigate the effects due to nonconservation of vorticity. The latter are reflected in the second term on the right hand side of Eq. (26). In a qualitative way, it is easy to see that these effects tend to *increase* the mean-square wind error E. Suppose, for example, that a new cyclone is generated in the hole. Since our method of analysis (which is based on the principle of vorticity conservation) cannot predict such developments, the predicted vorticity is too anticyclonic – i.e., $\nabla^2\varepsilon$ is negative. Thus, since ε vanishes at the boundary of the hole, ε is predominantly positive. This, coupled with the fact that S is positive for cyclogenesis, implies that the integral

$$\int_A \varepsilon S \, dA$$

is positive. The same conclusion, by exactly analogous reasoning, applies when $S < 0$ in A.

Since the second term on the right hand side of Eq. (26) is independent of time in the ensemble average, one can immediately deduce that E varies exponentially and approaches a positive limiting value determined by the condition that $\partial E/\partial t = 0$. That is,

$$\frac{\lambda C V E}{A} = \frac{2}{A}\int_A \varepsilon S \, dA \qquad (27)$$

Thus, if ε can be related to S, this formula provides a means of computing the limiting value of E. Not knowing very much about how S varies in detail, one would find it difficult to estimate the right hand side of Eq. (27) with any degree of exactness. On the other hand, it may be sufficient to estimate an upper bound on the limiting value of E; if the upper bound on E is substantially lower than the mean-square wind error of other

methods of analysis, the point is proved. Our next concern, therefore, is to set an upper bound on the right hand side of Eq. (27).

It will first be noted that, according to Schwarz' inequality,

$$|\overline{\varepsilon S}| \leq \sqrt{\overline{\varepsilon^2}}\sqrt{\overline{S^2}} \tag{28}$$

where the bar above a quantity now denotes its average over A. Next, we represent ε as an infinite series of the form

$$\varepsilon(x, y) = \sum_{n=1}^{\infty} a_n g_n(x, y) \tag{29}$$

in which the a_n's are constant and the functions $g_n(x, y)$ are defined by equations of the type

$$\nabla^2 g_n = -\alpha_n{}^2 g_n \text{ for } g_n = 0 \text{ on } C \tag{30}$$

Here the constants α_n are the eigenvalues for the equation and boundary conditions given above. We suppose that the terms of Eq. (29) are ordered in such a way that $\alpha_{n+1} > \alpha_n$. Now, if α_p and α_q are distinct eigenvalues and g_p and g_q are corresponding eigenfunctions, then Eq. (30) implies that

$$\int_A (g_p \nabla^2 g_q - g_q \nabla^2 g_p)\, dA = (\alpha_p^2 - \alpha_q^2) \int_A g_p g_q\, dA$$

or

$$\int_A \nabla \cdot (g_p \nabla g_q - g_q \nabla g_p)\, dA = (\alpha_p^2 - \alpha_q^2) \int_A g_p g_q\, dA$$

With the aid of Gauss' theorem, the left hand side of this equation may be transformed into

$$\oint_C \left(g_p \frac{\partial g_q}{\partial n} - g_q \frac{\partial g_p}{\partial n} \right) ds$$

Thus, since g_p and g_q vanish on C, the integral above vanishes and

$$\int_A g_p g_q\, dA = 0$$

This result shows that the functions g_p and g_q, corresponding to distinct eigenvalues α_p and α_q, are orthogonal in the region A.

Squaring both sides of Eq. (29) and noting that the functions $g_n(x, y)$ are orthogonal in A, we find that

$$\overline{\varepsilon^2} = \sum_{n=1}^{\infty} a_n^2 \overline{g_n^2} \tag{31}$$

Moreover, it follows from Eq. (30) that

$$\nabla^2 \varepsilon = -\sum_{n=1}^{\infty} a_n \alpha_n^2 g_n$$

whence, squaring both sides of this equation and integrating over A, we see that

$$\overline{(\nabla^2 \varepsilon)^2} = \sum_{n=1}^{\infty} a_n^2 \alpha_n^4 \overline{g_n^2}$$

$$= \alpha_1^4 \sum_{n=1}^{\infty} a_n^2 \left(\frac{\alpha_n}{\alpha_1} \right)^4 \overline{g_n^2}$$

Thus, since $\alpha_1 \leq \alpha_n$, the summation immediately above is term-for-term equal to or greater than the right hand side of Eq. (31), whence

$$\overline{\varepsilon^2} \leq \frac{1}{\alpha_1^4} \overline{(\nabla^2 \varepsilon)^2}$$

Combination of this result with the inequality (28) shows that

$$|\overline{\varepsilon S}| \leq \frac{1}{\alpha_1^2} \sqrt{\overline{S^2}}\sqrt{\overline{(\nabla^2 \varepsilon)^2}} \tag{32}$$

The error in vorticity may be estimated in the following way: Since we are seeking an upper bound on $\overline{\varepsilon S}$, we assume the most direct relationship between S and ε (or $\nabla^2 \varepsilon$). This, accordingly to Eq. (8), is

$$\frac{d}{dt} \nabla^2 \varepsilon \simeq -S$$

Moreover, integrating the equation above with respect to time, we see that

$$\nabla^2 \varepsilon(x, y) \simeq -T(x, y)\, \tilde{S}(x, y)$$

where $T(x, y)$ is the time required for a particle to traverse a trajectory ending at the point (x, y) and beginning at some point on C (where $\nabla^2 \varepsilon = 0$). The quantity $\tilde{S}(x, y)$ is the value of S averaged along such a trajectory. Introducing this result into the inequality (32), we have

$$|\overline{\varepsilon S}| \leq \frac{1}{\alpha^2} \sqrt{\overline{S^2}}\sqrt{\overline{T^2 \tilde{S}^2}}$$

Finally, since $\tilde{S}$ is an average of S, it is generally true that $\overline{\tilde{S}^2} \leq \overline{S^2}$. Thus, in view of the previous inequality,

$$|\overline{\varepsilon S}| \leq \frac{1}{\alpha_1^2}\sqrt{\overline{T^2}\,\overline{S^2}}$$

This inequality gives an upper bound on the value of the right hand side of Eq. (27).

As we have seen earlier, an upper bound on the right hand side of Eq. (27) corresponds to an upper bound on the limiting value of the mean-square wind error E, as given by Eq. (27). That is,

$$\frac{\lambda CVE}{A} = \frac{2}{\alpha_1^2}\sqrt{\overline{T^2}\,\overline{S^2}} \qquad (33)$$

This formula can be expected to yield an *over*-estimate of the limiting error E that can be attained by the dynamical methods outlined earlier. It will be recalled that the constant α_1 is the smallest eigenvalue for the region A; for a roughly circular region with average radius R, $\alpha_0 \simeq N/R$, where N is the smallest value of z for which $J_0(z) = 0$. If the region A is not very large, the quantity $\sqrt{\overline{T^2}}$ is about the average time required for individual particles to travel across the hole.

For the sake of concreteness, we now estimate an upper bound on the limiting value of the root-mean-square wind error for a circular hole. In this case, we take

$$\lambda = .80 \qquad C = 2\pi R$$

$$A = \pi R^2 \qquad \alpha_1 = \frac{2.4}{R}$$

Moreover, if the hole is not extremely large, $\sqrt{\overline{T^2}} \simeq R/V$. Approximately, then, Eq. (33) becomes

$$\sqrt{E} = .47\frac{R^2}{V}\sqrt{\overline{S^2}} \qquad (34)$$

Since V and $\sqrt{\overline{S^2}}$ probably do not vary rapidly with R in the ensemble average, this formula states that the limiting value of the root-mean-square wind error over a hole is roughly proportional to the square of the radius of the hole. In computing numerical estimates of the magnitude of the limiting root-mean-square wind error, we let $V = 600$ km/day, a value that has been found typical of holes 1,000 to 2,000 kms across. To estimate $\sqrt{\overline{S^2}}$, we first note that the mean absolute value of relative vorticity is of the order of 1/day. Thus, since numerical forecasts based on the principle of vorticity conservation correctly predict about 65 % of the variance of 24 hour changes, a typical value of $\sqrt{\overline{S^2}}$ is certainly less than 1/ (day)^2 and is probably about .5/(day)^2.

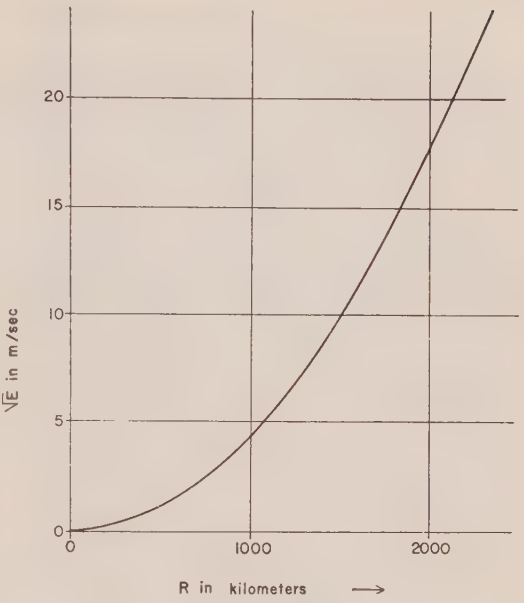


Fig. 5. Graph showing how the root-mean-square wind error $\sqrt{E}$ depends on the radius R of a hole.

The root-mean-square wind error $\sqrt{E}$ is plotted as a function of radius R on Fig. 5, for the values $V = 600$ km/day and $\sqrt{\overline{S^2}} = .5/$ (day)^2. Under the hypothetical but not unrealistic conditions we have assumed, Fig. 5 shows that the root-mean-square wind error over a hole 2,000 km across ($R = 1{,}000$ km) would be only about 5 m/sec. This value is not much larger than the combined errors of measurement and roundoff in standard wind observations, and is much less than the error of analysis carried out by the usual methods of interpolation and smoothing under comparable conditions. This conclusion applies strictly, of course, only if the true winds are known at all times and at all points outside the hole.[1] The results of further numerical experiments, designed to isolate the effects of frequency of observation and analysis error outside the hole, will be reported by Richardson.

[1] It should be emphasized, however, that our estimate of the limiting error is probably too large in this hypothetical case.

6. Summary and Conclusions

A procedure of repeated analysis and prediction, based on the dynamical principle of vorticity conservation, is proposed as a method of reconstructing meteorological conditions in a "hole" – i.e. a region from which no data are received, but which is surrounded by a region of frequent and dense observations. Mathematical analysis of the manner in which error is propagated by this method indicates that the mean-square wind error over a hole approaches a definite lower limit. In fact, if vorticity were conserved in actuality, the error would approach zero everywhere in the hole. Although this conclusion applies strictly only if the data are known at all times and at all points outside the hole, it has been substantiated by a series of numerical experiments carried out under conditions that simulate the true state of affairs.

It must be emphasized that the mean-square wind error over a hole will not (in general) approach a definite lower limit unless the vorticity is predicted by conserving vorticity along *forward*-extrapolated trajectories. Implausible as it may seem, this is the only method that guarantees that the correct vorticity from outside the hole is advected intact to points inside the hole.

In the general case, when vorticity is not conserved in actuality, the mean-square wind error over a hole still approaches a definite lower limit, which corresponds to a balance between the rates of growth and decay of error. The growth of error in a hole is due to the fact that vorticity is not really conserved; the rate of growth is roughly

$$\frac{2}{\alpha_1^2}\sqrt{\overline{T^2}}\,\overline{S^2}$$

where α_1 is the smallest value of α for which $\nabla^2 g = -\alpha^2 g$ and g vanishes on the boundary of the hole; $\sqrt{\overline{T^2}}$ is the root-mean-square value of the time elapsed since the particle at each point (x, y) in the hole last crossed the boundary of the hole; and $\overline{S^2}$ is the mean-square value of the individual rate of change of vorticity.

The decay of error in a hole is due to the advection of "correct" vorticity across the windward edge of the hole; the rate of decay is approximately

$$\frac{\lambda CVE}{A}$$

where λ is a constant, equal to about 0.8; C is the circumference of the hole; A is its area; V is the average absolute rate of flow across the boundary; and E is the mean-square wind error over the hole. The limiting value of E is obtained simply by equating the rates of growth and decay, as given above.

For typical values of V and $\overline{S^2}$, the asymptotic value of root-mean-square wind error over a hole 2,000 km in diameter is of the order of 5 m/sec and varies as the square of the diameter. This standard of accuracy is considerably higher than can be attained by conventional methods of analysis, based primarily on interpolation and smoothing.

The foregoing results strongly suggest that the most economical way of improving the quality of analysis over regions of low data density is not merely to set up more observing stations – particularly in ocean areas or other regions in which the establishment, maintenance and supply of observation posts is extremely expensive. In some such cases, it is probably more economical to improve the quality of analysis in fairly large holes by applying the data-processing methods studied here.

It should also be pointed out that these methods are not very effective unless the region along the windward edge of the hole is covered by enough observations to permit a fairly accurate estimate of the "incoming" vorticity. Accordingly, the optimum arrangement of fixed observing stations (with due regard to expense) is probably a weighted combination of:

(a) Uniform and fairly dense networks of stations over populated land areas and other areas where it is not so expensive to establish and maintain them. Over these regions, methods of pure interpolation and smoothing are probably adequate.

(b) Dense "strips" of stations, surrounding "holes" up to about 2,000 km in diameter, and with highest density along those edges of the holes which (on the average) lie to windward. Along these "strips", observations should be made frequently, in order to pro-

vide a continuous supply of "correct" vorticity for injection into the holes. (It has previously been pointed out that increasing the frequency of observation does not result in a proportionate increase in expense.) In the holes, of course, it would be necessary to apply dynamical methods of analysis. Although this course would probably require full-time use of a high-speed, large-capacity computing machine, the total cost of data-processing would be only a small fraction of the cost of the entire network of observing stations.

Acknowledgment

I am especially indebted to Captain Norman N. Richardson, who very kindly granted permission to reproduce some results of his numerical experiments in Figs. 4(a) and 4(b).

REFERENCES

Bergthorsson, P., and Döös, B., 1955: Numerical weather map analysis. *Tellus*, **7**, pp. 329—340.

Gilchrist, B., and Cressman, G. P., 1954: An experiment in objective analysis. *Tellus*, **6**, pp. 309—318.

Richardson, N. N.: *Licentiate thesis*, University of Stockholm, 1961. (Unpublished.)

The Effect of Advection in a Region of no Data

An alternative derivation of Thompson's Equation

By F. B. SMITH, Institute of Meteorology, Stockholm, and British Meteorological Office

(Manuscript received June 6, 1961)

In the preceding paper, Thompson has discussed the time variation of the mean square wind error in a hole (i.e. a region in which there are no data), surrounded by an area of complete and accurate data coverage. At the initial time the stream-function field inside the hole is estimated by some objective method consistent with its known distribution on the boundary between the hole and the complete-data exterior. This estimate will differ from the true field and both the vorticity and advection wind velocity are locally in error within the hole and on the boundary. The question arises, can this error be significantly reduced as time proceeds by taking advantage of the advection of good data into the hole from outside? Noting that both the erroneous advection velocity and the possibility of cyclogenesis within the hole (not predictable by a barotropic scheme) may tend to work against the desired reduction, Thompson has shown that the error tends to approach some balance position determined by the size of the hole, the average wind speed and the cyclogenesis. Richardson in the succeeding paper verifies these general conclusions and shows that in general the subsequent average error is significantly better than that obtained by objective guesswork.

In this note an alternative derivation of Thompson's result is given. As shown by Richardson, the derivation reduces to a particularly simple form when the wind velocity can be assumed constant over the hole, and the cyclogenesis neglected. In this case the wind error approaches quasi-exponentially to zero.

The general derivation follows the following lines. The time variation of the mean square vector wind error over the hole is considered. It is noted, following Thompson, that the error part of the advection velocity field plays no part in this variation, it only redistributes error *within* the hole. The true advection velocity and the estimated velocity field can therefore, for the purposes of advection, be taken as equal. The rate of error change then depends on an area integral over the hole, representing the Lagrangian advection of the error field (specifically involving the error vorticity), and on an area integral of the cyclogenetic effect. Since the error in the velocity is discontinuous on the boundary, the error vorticity is singular there, and the advection integral may be considered in three parts, one coming from the boundary representing the advection of good data into the hole along the windward edge, a second from the interior which represents the net advection of error *out* of the hole (calculated as if the error field varied continuously across the boundary), and thirdly a term introduced by variations of velocity within the hole. This last term involves the average correlation between the true and the error velocity fields and acts *in this context* as a spurious cyclogenetic effect. In an ensemble average over many atmospheric situations this term must be zero, and even in individual cases the results presented by Richardson for his purely barotropic experiments indicate that, in those circumstances at least, it is negligibly small.

Therefore to a very good approximation the rate of change of the mean square wind velocity error depends on the difference of the real cyclogenetic term and the net outflux of error from the hole (including the boundary). The

final result is identical to that obtained by Thompson.

Define ε = local error in the stream function. Then the mean square wind error is

$$E = \frac{1}{A}\int_A \nabla\varepsilon \cdot \nabla\varepsilon \, dA$$

where A is the area of the hole, boundary C. Let C_i be the part of C with inflowing velocities, C_o the part with outflowing velocities. Let V be the local velocity, V_n the *inward* normal component of V on C, and S the Lagrangian rate of change of vorticity of a fluid element.

Thompson has shown that the error component of the advection velocity plays no part in changing E, and therefore

$$\frac{\partial E}{\partial t} = \frac{2}{A}\int_A \varepsilon\left(\frac{d}{dt}\nabla^2\varepsilon - \frac{\partial}{\partial t}\nabla^2\varepsilon\right) dA + \frac{2}{A}\int_A \varepsilon S dA$$

(his equation (10)), in which the advection velocity is considered to be the true local velocity.

Let R_1 be the contribution to the first of these integrals from the boundary where $\nabla^2\varepsilon$ has a singularity. Since the error field is zero outside A, the integral may be considered to extend over all space. We consider a vanishingly narrow strip centred on C, width 2δ. Now

$$\varepsilon\left(\frac{d}{dt}\nabla^2\varepsilon - \frac{\partial}{\partial t}\nabla^2\varepsilon\right) =$$

$$= \lim_{\Delta t\to 0}\frac{1}{\Delta t}\varepsilon(\mathbf{r})\left[\nabla^2\varepsilon(\mathbf{r}) - \nabla^2\varepsilon(\mathbf{r} - \mathbf{V}\Delta t)\right]$$

where $\mathbf{r}$ is a general position vector.
Since $\varepsilon(\mathbf{r})$ is continuous across C, the only contribution comes from the singularity in each of the two $\nabla^2\varepsilon$, the first when $\mathbf{r}$ lies within the strip, the second when $\mathbf{r} - \mathbf{V}\Delta t$ lies in the strip. Then, if $\mathbf{s}$ is a general position vector on C, and $\mathbf{n}$ is a unit inward-pointing vector:

$$R_1 = \frac{2}{A}\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_C\int_{s-\delta n}^{s+\delta n}\left[\varepsilon(\mathbf{p})\nabla^2\varepsilon(\mathbf{p}) - \right.$$
$$\left. - \varepsilon(\mathbf{p} + \mathbf{V}\Delta t)\nabla^2\varepsilon(\mathbf{p})\right] d\mathbf{p}$$

Now, on C_o, $\varepsilon(\mathbf{p} + \mathbf{V}\Delta t)$ lies outside A. Therefore $\varepsilon(\mathbf{s}) - \varepsilon(\mathbf{s} + \mathbf{V}\Delta t) \equiv 0$ and there is no contribution to R_1. Physically this states that good data (zero error) is not advected into the hole on the outflow edge.
Since $\varepsilon(\mathbf{p})$ is continuous

$$R_1 =$$
$$= \frac{2}{A}\lim_{\Delta t\to 0}\int_{C_i}\left[\frac{\varepsilon(\mathbf{s}) - \varepsilon(\mathbf{s} + \mathbf{V}\Delta t)}{\Delta t}\right]\int_{s-\delta n}^{s+\delta n}\nabla^2\varepsilon(\mathbf{p})d\mathbf{p}$$

Expanding

$$\varepsilon(\mathbf{s} + \mathbf{V}\Delta t) = \varepsilon(\mathbf{s}) + \mathbf{V}_n\Delta t\frac{\partial\varepsilon}{\partial n}(s+)$$

where terms of higher power in Δt have been omitted because in the limit they vanish. The property that the gradient of ε along the boundary is zero has been used, $\frac{\partial\varepsilon}{\partial n}(s+)$ refers to the normal gradient on the interior side of C_i.

Similarly $\int_{s-\delta n}^{s+\delta n}\nabla^2\varepsilon(\mathbf{p})\,d\mathbf{p} = \frac{\partial\varepsilon}{\partial n}(s+)$

Therefore $R_1 = -\frac{2}{A}\int_{C_i} V_n\nabla\varepsilon\cdot\nabla\varepsilon\, d\mathbf{s}$

R_1 has accounted for all the discontinuity on C. The remaining contribution $(R_2 + R_3)$ to the first integral in $\frac{\partial E}{\partial t}$ must therefore come from the interior of A and, in so far as the interior comes right up to the boundary C, no contribution from the discontinuity must be allowed to enter these terms. This is achieved by assuming there exists an error field which proceeds *continuously* out from A into the exterior.

$$R_2 + R_3 = \frac{2}{A}\int_A \varepsilon\left(\frac{d}{dt}\nabla^2\varepsilon - \frac{\partial}{\partial t}\nabla^2\varepsilon\right)dA$$
$$= \frac{2}{A}\int_A \varepsilon\left(\frac{d}{dt}\nabla^2\varepsilon - \nabla\cdot\frac{d}{dt}\nabla\varepsilon\right)dA +$$

$$\frac{2}{A}\int_A \nabla\cdot\varepsilon\left(\frac{d}{dt}\nabla\varepsilon-\frac{\partial}{\partial t}\nabla\varepsilon\right)dA-$$

$$\frac{2}{A}\int_A \nabla\varepsilon\cdot\left(\frac{d}{dt}\nabla\varepsilon-\frac{\partial}{\partial t}\nabla\varepsilon\right)dA$$

Call R_2 the first term, and R_3 the last two. R_2 arises from the variations of $\mathbf{V}$ within A, and we can show that on average this term is zero. In Richardson's theory, where $\mathbf{V}$ is constant, it is identically zero. Since on any one occasion this term cannot be predicted it can be termed a spurious cyclogenetic effect:

$$R_2=\frac{2}{A}\int_A \varepsilon S'\,dA$$

where

$$S'=(\mathbf{V}\cdot\nabla)(\nabla\cdot\nabla\varepsilon)-\nabla\cdot(\mathbf{V}\cdot\nabla)\nabla\varepsilon$$

Expressing $\mathbf{V}$ in terms of a stream-function ψ, S' can be written as

$$S'=\frac{\partial^2\varphi}{\partial x\partial y}\left[\frac{\partial^2\varepsilon}{\partial x^2}-\frac{\partial^2\varepsilon}{\partial y^2}\right]-\frac{\partial^2\varepsilon}{\partial x\partial y}\left[\frac{\partial^2\varphi}{\partial x^2}-\frac{\partial^2\psi}{\partial y^2}\right]$$

Using the same technique employed by Thompson we can express the second term in a coordinate system (X, Y) at 45° to (x,y). Then

$$S'=\frac{\partial^2\psi}{\partial x\partial y}\left[\frac{\partial^2\varepsilon}{\partial x^2}-\frac{\partial^2\varepsilon}{\partial y^2}\right]-\frac{\partial^2\varphi}{\partial X\partial Y}\left[\frac{\partial^2\varepsilon}{\partial X^2}-\frac{\partial^2\varepsilon}{\partial Y^2}\right]$$

Since in the average there is no preferred direction for the perturbation fields these two similar terms are statistically equal and $\overline{S'}\to 0$. For individual cases R_2 may be incorporated with the real cyclogenetic term to form a term with an apparent cyclogenetic factor $S^\star = S + S'$.

Returning to R_3, we note that the first of the two integrals disappears, since applying the divergence theorem it is equal to the line integral of a factor proportional to ε around C (where ε is zero). This use of the divergence theorem is valid for the continuous field. Thus:

$$R_3=-\frac{1}{A}\int_A\left[\frac{d}{dt}(\nabla\varepsilon\cdot\nabla\varepsilon)-\frac{\partial}{\partial t}(\nabla\varepsilon\cdot\nabla\varepsilon)\right]dA$$

$$=-\frac{1}{A}\int_A \mathbf{V}\cdot\nabla(\nabla\varepsilon\cdot\nabla\varepsilon)\,dA$$

$$=-\frac{1}{A}\int_A \nabla\cdot(\nabla\varepsilon\cdot\nabla\varepsilon)\,\mathbf{V}\,dA$$

since $\mathbf{V}$ is non-divergent.

Applying the divergence theorem for the assumed continuous field

$$R_3=\frac{1}{A}\int_C V_n\,\nabla\varepsilon\cdot\nabla\varepsilon\,d\mathbf{s}$$

where V_n is the inward normal component. This term represents the net advection of the error-field out of A. Substituting R_1, R_2 and R_3 back into $\partial E/\partial t$, we obtain

$$\frac{\partial E}{\partial t}=-\frac{1}{A}\int_C|V_n|\,\nabla\varepsilon\cdot\nabla\varepsilon\,d\mathbf{s}+\frac{2}{A}\int_A\varepsilon S^\star\,dA$$

This is basically identical to Thompson's result (equation (26)).

The author is indebted to the Director-General, British Meteorological Office, for permission to publish this note.

Numerical Tests of a Method for Dynamic Analysis in Regions of Poor Data Coverage[1]

By NORMAN N. RICHARDSON, Captain, USAF,
International Meteorological Institute, Royal University of Stockholm

(Manuscript received May 10, 1961)

Abstract

This paper describes a series of numerical experiments which were done to test a theoretical method for dynamic analysis in regions of poor data coverage. Theoretically, if vorticity were conserved, the exact vorticity field in such a region could be determined by Lagrangian advection of vorticity from areas of known data. The technique was applied to a field of rather dense data where a "hole" was created by withholding all observations within a given area.

The first experiments, assuming the atmosphere to be purely barotropic, resulted in a field within the hole that closely approached, after a reasonable time, the field obtained by a barotropic forecast using *all* the data. The remaining experiments, under simulated operational conditions, tested the effects of: 1) baroclinic processes, 2) size of hole, 3) frequency of observation, and 4) allowing a single strip of data across the hole. The results were much as predicted by the theory and reasonably accurate analyses were obtained.

1. Introduction

For many years meteorologists frequently blamed inaccuracies in upper air analysis on "lack of data". Since World War II, however, it has become increasingly apparent that actually there exists too much data to be readily assimilated and evaluated by individual or small groups of forecasters. The development of high speed electronic computers has, to a large extent, overcome this difficulty. We are again faced with "lack of data"-problems in many large areas of the world.

These areas with a low density of observations are distinct handicaps in any large scale forecasting technique because they prevent an accurate analysis of the existing meteorological situation. All synoptic, kinematic, dynamic and most statistical forecasting techniques are based on these analyses. The dynamic techniques are particularly sensitive to inaccurate analysis due to the mechanics of the method. A forecast at a given point depends not only on the data upstream but also on an ever widening band of data surrounding the upstream trajectory. This "region of dependence" is explicitly determined by the character of the equations used in the forecasting procedure, and may be, in some cases, extremely large. It is obvious, therefore, that forecasts for areas downstream from regions of low density of observations will be greatly affected by inaccuracies in the analysis.

P. D. Thompson (1961) reasoned that for all the inaccurate data "flowing *out*" of these low density observation areas, an almost equal amount of accurate data was "flowing *in*" on the other side. It remained thus to develop a

[1] The research reported here has been sponsered in part by the Geophysics Research Directorate, AFCRL, AFRD, of the Air Research and Development Command, USAF, through its European Office.

technique for utilizing these accurate data in determining a more correct analysis in these regions. Thompson has developed a theory which showed that a more correct analysis could be obtained by conserving vorticity along forward extrapolated trajectories. In effect, the accurate data advects into the region and, to the extent that vorticity is conserved in actuality, replace the inaccurate data formerly there. He has also discussed the effect of baroclinicity, *i.e.* the creation of vorticity within the region. This effect is not well understood but its influence cannot be avoided.

Conventional barotropic forecasts are fairly accurate in regions of good data coverage. RMS vector wind errors average about 7 m.p.s. in 24 hour forecasts. In areas of poor data coverage, however, they have not been as accurate and in 24 hour forecasts, RMS vector wind errors average about 10 m.p.s.

This paper will describe a set of numerical experiments, using a barotropic model, which were done to test Thompson's theory.

2. Theory

Thompson has shown that the reduction of error in a sparse data region depends on: (1) the size of the region, (2) the speed of advection (this determines the amount of accurate data flowing into the region), and (3) the magnitude of the initial error. The sparse data region is termed a "hole" which, in theory, contains *no* observations. He further shows that when cyclogenesis is ignored, the dominant effect is due to advection of the accurate field by the *mean* wind **V** on the boundary of the hole. The contributions arising from deviations from **V** are of secondary importance in the advection and will be ignored in the following simplified theory.

General Definitions

A = area of "hole"
$\mathbf{r}$ = position vector inside A
C = boundary of A
ψ = stream function
$\zeta = \nabla^2\psi$

We assume that we know the distribution of the streamfunction ψ everywhere at all times outside the hole. Inside the hole, we approximate (by guess or "educated" guess depending on our knowledge of conditions surrounding the hole) a distribution of streamfunction ψ_a. We now define the mean error in the hole:

$$E = \frac{1}{A}\int_A (\nabla\varepsilon\cdot\nabla\varepsilon)\,dA \tag{1}$$

where ε is the error in our approximate streamfunction $\psi_a - \psi$. Differentiating with respect to time, we obtain:

$$\frac{\partial E}{\partial t} = \frac{1}{A}\int_A \frac{\partial}{\partial t}(\nabla\varepsilon\cdot\nabla\varepsilon)\,dA \tag{2}$$

Since ε is 0 everywhere outside the hole (and on the boundary), but is generally finite inside the hole, $\nabla\varepsilon$ will be discontinuous across the boundary C of A. Equation (2), therefore, will have contributions from the interior of A and from the boundary C. We shall solve for the boundary contribution first. Equation (2) may be transformed to:

$$\frac{\partial E}{\partial t} = \frac{2}{A}\int_A \nabla\varepsilon\cdot\nabla\frac{\partial\varepsilon}{\partial t}\,dA = \frac{2}{A}\int_A \nabla\cdot\varepsilon\nabla\frac{\partial\varepsilon}{\partial t}\,dA - \frac{2}{A}\int_A \varepsilon\nabla^2\frac{\partial\varepsilon}{\partial t}\,dA \tag{3}$$

Noting the form of the first integral, we transform it to a line integral by Gauss' Theorem. The integrand then vanishes because ε is 0 on the boundary C. The second integral contains the term $\frac{\partial}{\partial t}\nabla^2\varepsilon$. Remembering that vorticity (defined as the Laplacian of a streamfunction) is conserved, we recognize this term as the local change of vorticity error which is merely the advection of $\nabla^2\varepsilon$ by the mean wind **V**. Thus:

$$\frac{\partial E}{\partial t} = \frac{2}{A}\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_A \varepsilon(\mathbf{r})\left[\nabla^2\varepsilon(\mathbf{r}) - \nabla^2\varepsilon(\mathbf{r}-\mathbf{V}\Delta t)\right]d\mathbf{r} \tag{4}$$

Considering the contribution of this integral at time t, we see that advection during a time Δt previous to time t crosses the boundary only on the inflow side. Therefore we may restrict our attention, in considering the boundary effect, to that part of the boundary

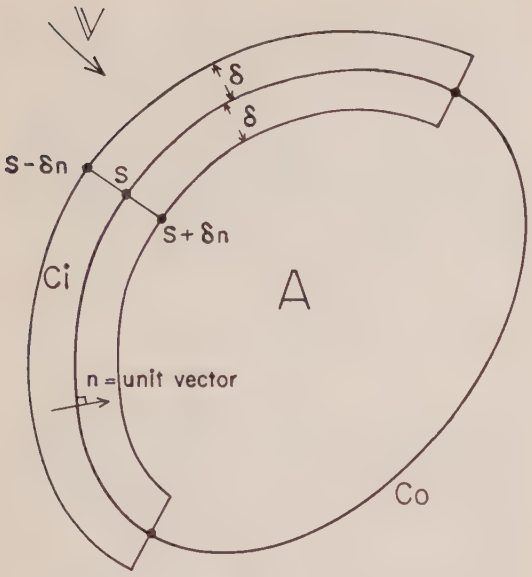


Fig. 1.

Ci across which there is flow into the hole (see Fig. 1).

This contribution comes from a strip along Ci whose width ($2\,\delta$) is vanishingly small. It is the behaviour of $\nabla^2\varepsilon$ on Ci that determines this contribution. Therefore, the first term in the integrand of (4) makes a contribution when $\mathbf{r}$ is integrated over the strip and the second term when $(\mathbf{r}-\mathbf{V}\,\Delta t)$ is integrated over the strip. Combining these two integrals and using ϱ as a position vector of integration, we obtain:

$$\left[\frac{\partial E}{\partial t}\right]_c = \frac{2}{A}\lim_{\Delta t\to 0}\frac{1}{\Delta t}\left\{\lim_{\delta\to 0}\int_{Ci}\int_{\mathbf{S}-\delta\mathbf{n}}^{\mathbf{S}+\delta\mathbf{n}}[\varepsilon(\varrho)\,\nabla^2\varepsilon(\varrho) - \varepsilon(\varrho+\mathbf{V}\,\Delta t)\,\nabla^2\varepsilon(\varrho)]\,d\varrho\right\} \qquad (5)$$

where $\mathbf{n}$ is the inward-directed unit vector and $\mathbf{S}$ is the general position vector on Ci. Performing the inner integration and noting that $\frac{\partial\varepsilon}{\partial n}=0$ in the outward direction, we obtain:

$$\left[\frac{\partial E}{\partial t}\right]_c = \frac{2}{A}\lim_{\Delta t\to 0}\frac{1}{\Delta t}\int_{Ci}\left[\varepsilon(\mathbf{S})\frac{\partial\varepsilon}{\partial n}(S+) - \varepsilon(S+\mathbf{V}\,\Delta t)\frac{\partial\varepsilon}{\partial n}(S+)\right]ds \qquad (6)$$

Taylor expanding the function $\varepsilon(S+\mathbf{V}\Delta t)$ about the point $\varepsilon(s)$ we obtain:

$$\varepsilon(S+\mathbf{V}\,\Delta t)=\varepsilon(S)+\mathbf{V}_i\,\Delta t\frac{\partial\varepsilon}{\partial n}(S+) \qquad \text{(for small } t\text{)}$$

where $\mathbf{V}_i$ is the component of $\mathbf{V}$ parallel to $\mathbf{n}$. Therefore:

$$\left[\frac{\partial E}{\partial t}\right]_c = -\frac{2}{A}\int_{Ci}\mathbf{V}_i\frac{\partial\varepsilon}{\partial n}\cdot\frac{\partial\varepsilon}{\partial n}\,ds = -\frac{2}{A}\int_{C_i}\mathbf{V}_i\,\nabla\varepsilon\cdot\nabla\varepsilon\,ds \qquad (7)$$

The contribution from the interior of A must now be considered. Discontinuities in $\nabla\varepsilon$ on C can now be ignored. Equation (2) can be written:

$$\left[\frac{\partial E}{\partial t}\right]_A = -\frac{1}{A}\int_A\mathbf{V}\cdot\nabla(\nabla\varepsilon\cdot\nabla\varepsilon)\,dA = -\frac{1}{A}\int_A\nabla\cdot(\nabla\varepsilon\cdot\nabla\varepsilon)\,\mathbf{V}\,dA \qquad (8)$$

Since $\mathbf{V}$ is non-divergent

$$\left[\frac{\delta E}{\delta t}\right]_A = +\frac{1}{A}\oint_c\mathbf{V}_n\,\nabla\varepsilon\cdot\nabla\varepsilon\,dS \qquad (9)$$

where $\mathbf{V}_n$ is the inward normal component of $\mathbf{V}$. Combining (7) and (9) we find that the total rate of change of E is:

$$\frac{\partial E}{\partial t} = -\frac{1}{A}\left\{\int_{Ci}\mathbf{V}_i\,\nabla\varepsilon\cdot\nabla\varepsilon\,dS+\int_{C_0}\mathbf{V}_0\,\nabla\varepsilon\cdot\nabla\varepsilon\,ds\right\} \qquad (10)$$

where $\mathbf{V}_0=-\mathbf{V}_n$ on C_0 (see fig. 1).

Approximately $\frac{\partial E}{\partial t}=-\frac{1}{A}C\,\overline{|\mathbf{V}_n|}\,\overline{(\nabla\varepsilon\cdot\nabla\varepsilon)}$

where the bar denotes an average around C, therefore:

$$\frac{\partial E}{\partial t}=-\frac{\lambda C\,\overline{|\mathbf{V}_n|}\,E}{A} \qquad (11)$$

where λ is the ratio of the value of $\overline{\nabla\varepsilon\cdot\nabla\varepsilon}$ over the boundary to its value over A, and is assumed to be approximately constant.

3. Numerical Experiments

The numerical experiments were designed to serve two distinct purposes. *First*, the general validity of the theory must be checked. *Second*, a scheme that could be used in actual practice must be tested and evaluated. The first set of experiments, therefore, was performed without considering the application of the method on an operational basis. The second set, however, fulfilled the requirements that would have to be met in actual forecasting operations.

A. *Testing the Theory*

The simplest method for testing the theory appeared to be some technique whereby vorticity was actually conserved. It was decided, therefore, to use the one-parameter barotropic model and to assume that its results were perfect and represented the "true" state of the atmosphere. These "barotropic forecasts" provided "true" data *outside* the hole, which were used in making the "hole forecast", and "true" data *inside* the hole, which were used in evaluating the results of the "hole forecast".

Initially, we knew the observed distribution of ψ everywhere and could compute $\zeta = \nabla^2\psi$ everywhere. This information, with a suitable set of boundary conditions, was all we needed for the barotropic forecast. For the hole forecast however, we assumed that we knew nothing inside the hole and must depend only on initial observations *outside* the hole and subsequent "true" barotropic forecast data *outside* the hole. In our first computations, we guessed a value of relative vorticity $\zeta_a = 0$ at all points within the hole. This guessed field was then smoothed to eliminate discontinuities at the boundary of the hole. This smoothing was done by the following operator:

$$\zeta_a\,(\text{smoothed}) = \frac{1}{4+a}\,(\zeta_1 + \zeta_2 + \zeta_3 + \zeta_4 + a\,\zeta_0) \tag{12}$$

where the subscripts denote the positions of the gridpoints and a is a weighting factor applied to the center point. This equation may also be written:

$$\zeta_a\,(\text{smoothed}) = \frac{1}{4+a}\,(\zeta_1 + \zeta_2 + \zeta_3 + \zeta_4 + \\ + (4+a)\,\zeta_0 - 4\zeta_0) = \zeta_0 + \frac{1}{4+a}\,\nabla^2\zeta \tag{13}$$

This simplifies the computing procedures. A series of hand computations were made with this operator using values of 1, 2, and 3 for n. Three iterations were made. Based on the results of these computations, n was given the value 2 and 3 iterations were performed in the numerical experiments. In later computations, the approximate vorticity in the hole was computed from the known circulation on the boundary of the hole.

$$\overline{\zeta}_a = \frac{1}{A}\oint \mathbf{V}s\,ds \tag{14}$$

The smoothing operator was also applied to this field, but due to the limitations of our technique, no check was made to determine that the smoothed field agreed with the circulation.

We then had an "accurate" ψ-field and ζ-field *outside* the hole and an approximate ζ_a-field *inside* the hole. Using the "accurate" ψ-values on the boundary of the hole, we solved the Poisson equation $\nabla^2\psi = \zeta_a$ by relaxation and obtained the ψ_a-field inside the hole. This was all the information needed to compute the advection of vorticity inside the hole.

The advection was computed in a quasi-Lagrangian way. The components of velocity were computed at each grid point in the hole from the ψ_a-field. Assuming that these components represent a mean velocity of advection during the 45 minute time step, they determined the origin of a trajectory which terminated at the grid point at the end of the time step. The forecast vorticity at time t at the grid point, therefore, will be the same as that at the origin at time $(t - 45)$ min. The absolute vorticity at this origin was computed by linear extrapolation from the vorticity values at the adjacent grid points.

We then computed a 45 minute forecast over the entire region, using all the observed data, with the barotropic model. This gave us a forecast ψ-field and ζ-field everywhere. We again relaxed the ζ_a-field inside the hole (using the "accurate" forecast ψ on the boundary) to get the ψ_a forecast field. We then had all the required information to repeat the advection computation inside the hole. This procedure was repeated until the desired length of forecast was obtained. Print-outs of both the "accurate" barotropic fields and the

hole data were made at periodic intervals. The error was evaluated from these print-outs.

The above procedure is summarized in the step-by-step outline below:

1. Determine the initial ζ-field from the observed ψ-field over the entire region.
2. Approximate an initial ζ_a-field inside the hole.
3. Smooth this approximate field to eliminate boundary discontinuities.
4. Determine an approximate ψ_a-field inside the hole from the approximate ζ_a-field and the known ψ-field on the boundary of the hole.
5. Compute the forecast ζ_a-field inside the hole by the quasi-Lagrangian advection method.
6. Compute the forecast ψ-field and ζ-field with the barotropic forecast.
7. Return to step 4 and use the new forecast values.

B. *Operational Computations*

The second set of experiments was designed to fulfill operational requirements; *i.e.* provision was made to utilize *new observed* data outside the hole as it became available. Also, the actual known data within the hole was never used at any time during the computations.

Initially, we used only the known ψ-field outside the hole. As before, we computed the vorticity ζ from this field. The approximate vorticity field ζ_a inside the hole was again guessed or computed from the circulation, and these fields were "faired" at the boundary by using the smoothing operator. The ψ_a-field inside the hole was determined, as before, by relaxing the ζ_a-field and using known ψ-values on the boundary. Note that the vorticity values on the boundary of the hole were determined only from the initial approximate values in the hole with some influence (during the smoothing) coming from the known vorticity along the first grid row outside the hole. This did not guarantee that the boundary vorticity values would be in agreement with the ψ-field. Therefore, we computed a new ζ'-field from the combined ψ- and ψ_a-fields. This did not change the vorticity values outside the boundary of the hole or within the hole, only those along the boundary. These new values on the boundary then agreed with the ψ-distribution and we proceeded with the quasi-Lagrangian advection as before.

When the computations reached a new observation time, all the data were discarded except the forecast ζ_a-field inside the hole. This, as before, was used with the new observed ψ-values on the boundary of the hole to determine the ψ_a-field within the hole. The new ζ'-field was computed from the combined ψ- and ψ_a-fields and we proceeded to a new observation time.

The above procedure is summarized in the step-by-step outline below:

1. Determine the initial ζ-field outside the hole from the observed ψ-field outside the hole.
2. Approximate an initial ζ_a-field inside the hole.
3. Smooth this approximate field to eliminate boundary discontinuities.
4. Determine an approximate ψ_a-field inside the hole from the approximate ζ_a-field and the known ψ-values on the boundary of the hole.
5. Compute an "adjusted" ζ'-field from the combined ψ- and ψ_a-fields (to insure compatibility).
6. Compute a forecast ζ-field by quasi-Lagrangian advection over the entire field.
7. Relax the forecast ζ-field to obtain a forecast ψ-field over the entire region.
8. Return to step 6 and repeat until a new observation time is reached.
9. At the new observation time, retain *only* the ζ_a-field inside the hole, return to step 4 and repeat.

4. Results

A. The first set of experiments, to test the theory, was based on the observed data at 1200Z on 25 April 1960. The hole was chosen over the North Sea, Northern Germany, and Southern Scandinavia. A larger hole was subsequently chosen in the same area. This particular set of data was chosen because the one-parameter model barotropic forecasts were particularly accurate during this period. The hole locations were chosen to have a relatively dense network of observations around the boundaries. A section of the 16×16 grid,

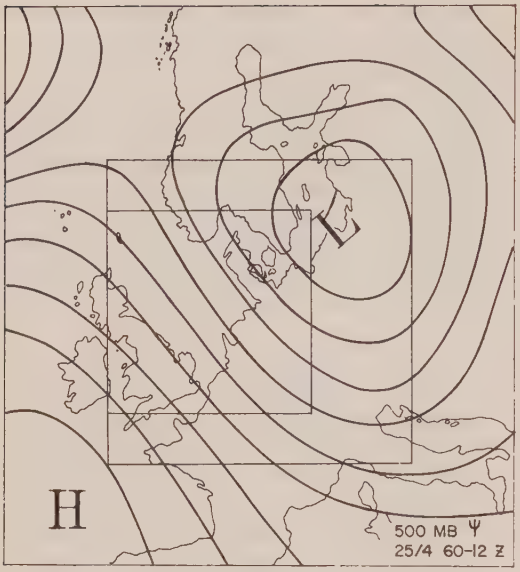


Fig. 2.

the location of the holes, and the initial synoptic pattern are shown in figure 2.

Early computations in the 3×3 hole (1200×1200 km) were encouraging in that

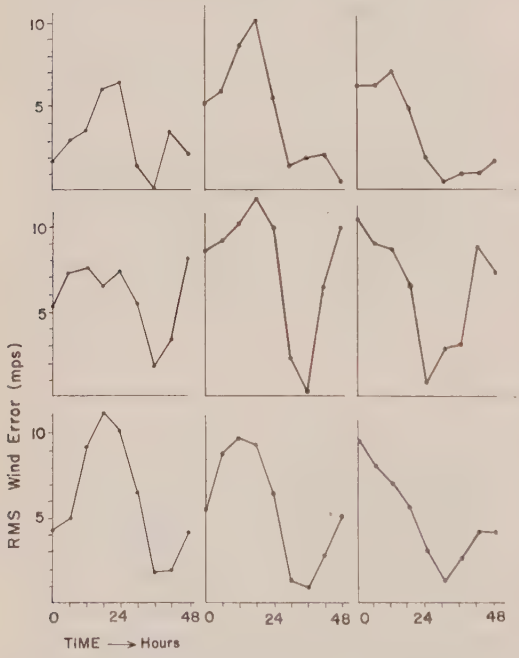


Fig. 3. The change of RMS wind error with time at each grid point of a 3×3 hole. The wind flow is from upper right to lower left.

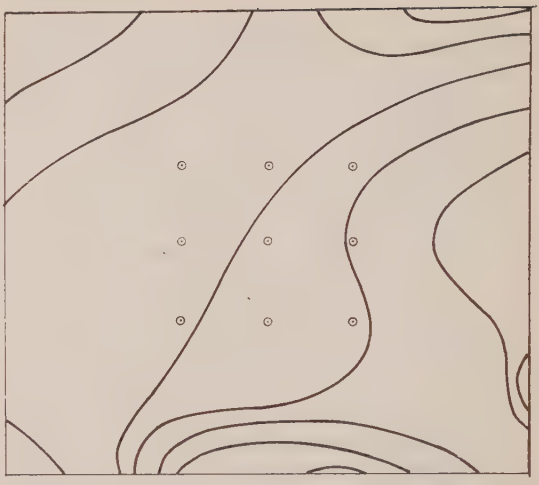


Fig. 4. The absolute vorticity field (as computed by quasi-Lagrangian advection) after 5 hours. Grid points in the hole are circled.

the decrease of error advected across the hole just as predicted in the theory (see fig. 3). The decrease occurred first at the points along the inflow boundary and the points along the outflow boundary showed the decrease last. At most points, however, this decrease was only temporary and the error grew rapidly almost immediately after having decreased to a very small value. These results were very similar to those experienced by WELANDER (1956) in a similar experiment.

A careful analysis was made of all the barotropic fields and the hole forecast fields to find an explanation of this growth of error. It was immediately apparent that the fields generated by the quasi-Lagrangian method were very smooth and all the small scale perturbations had been smoothed out. On the other hand, the fields generated by the barotropic forecast (using an Eulerian scheme), were definitely not smooth and many small scale perturbations were present (see figs. 4 & 5). Remembering that these fields must match at the boundary of the hole, it was apparent that truncation errors alone could cause the technique to fail. In some cases, the uneven field of barotropic data resulted in an erroneous change of sign during the advection.

The theory required that the vorticity be conserved along *forward extrapolated trajectories*. Therefore, the only alternative was to modify the barotropic forecast routine to use the

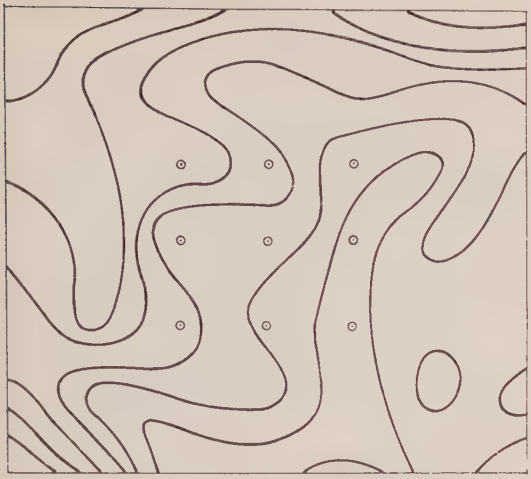


Fig. 5. The barotropic Eulerian absolute vorticity field after 6 hours.

type of advection rather than the usual Eulerian. This proved to be the correct solution (see fig. 6). The decrease in error still progressed downstream, it approached zero after a reasonable length of time, and there was no tendency for later growth beyond that which could be due to round-off error.

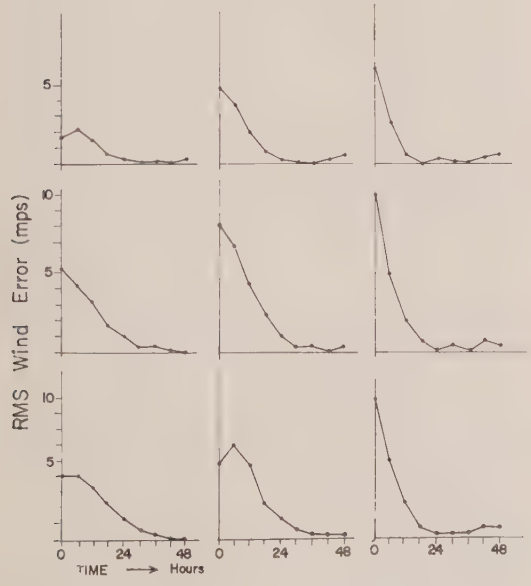


Fig. 6. The change of RMS wind error with time at each grid point of a 3×3 hole based on a quasi-Lagrangian forecast. The wind flow is from upper right to lower left.

Figures 7—9 show a comparison of the theoretical decrease of error (according to equation 11) and the observed decrease of error. The constant λ was evaluated from the data shown in figure 7, and then used to determine the theoretical curves in figures 8 and 9. The case shown by figure 9 differs from that shown by figure 8 *only* in the method of determining the initial approximate ζ_a-field in the hole. In the case shown by fig. 8, the initial relative vorticity was assumed to be zero. The other case was computed with an initial mean vorticity computed from the

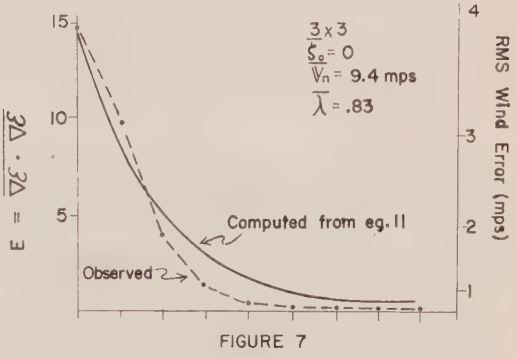


FIGURE 7

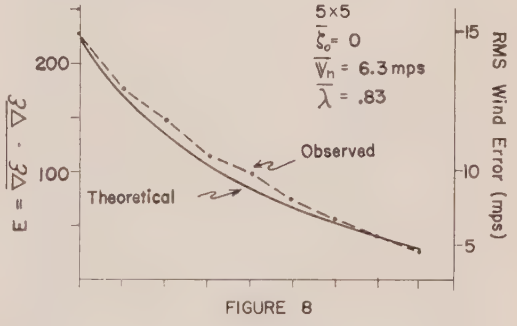


FIGURE 8

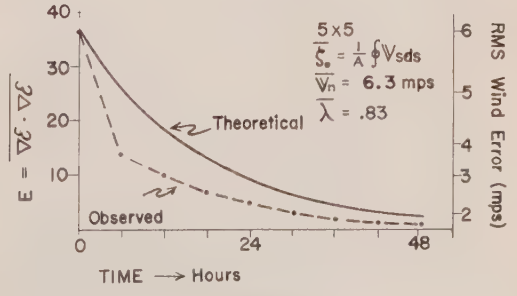


FIGURE 9

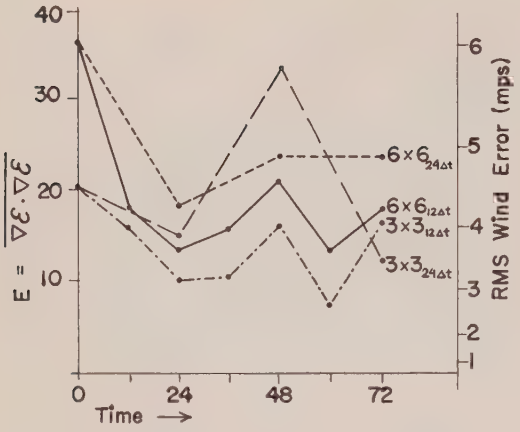


Fig. 10. Change of E with time for two hole sizes and two observation frequencies.

circulation around the hole. Note the *initial* RMS wind error was reduced from approximately 15 mps to 6 mps by this refinement only. Also, the latter case showed an even greater decrease of error than the theoretical one throughout the forecast period.

B. The second set of experiments, to simulate actual operational use, was based on the observed data for the period 0000Z 14 November 1960 to 1200Z 16 November 1960. As before, this period was chosen because the barotropic forecast had been quite successful during this time. Two holes, 3×3 grid points (1200×1200 km) and 6×6 grid points (2100×2100 km) were chosen in the same region as before. The total field consisted of 30×30 grid points.

There were two important differences between these and the experiments described earlier. First, as stated in section III b, new observed data were introduced into the scheme at the international observation times. Second, the results of the hole forecast were compared with actual observed data to compute the error field. Therefore, we did not expect the error to decrease to zero as vorticity is not actually conserved in the atmosphere. Rather, we expected the error to decrease to and oscillate about some minimum value.

In addition to varying the hole size, we also varied the time interval between the introduction of new observed data. This would give us an idea of the effect of hole size and frequency of observation on the method. In all, four complete computations were made, each for a period of 72 hours.

Figure 10 shows the results of these computations. As expected, the 12 hour observation frequency produced better results, both in amount and rate of error decrease. The effect of the hole size did not behave exactly as expected. In fact, the percentage decrease of error was slightly larger in the case of the larger hole and also, the rate of decrease was slightly greater. This does not really contradict the theory however, for the initial guess of vorticity (based on the circulation) would be much better in the small hole. Had we started the computations with zero vorticity in both cases, the smaller hole case would probably have been better.

The unexpectedly large error (larger than the initial value) at 48 hours in the $3 \times 3 - 24$ hour time interval case was not completely understood. There was a general increase of error between 24 and 48 hours in all cases, but it was excessively large in only the one instance. Possibly the natural oscillation of the error and some other effects (due to the long period of time between observations) both acted "in phase" to create a resonance. Also, this particular case may have possessed some unusual and unknown characteristics which made it misbehave. Although identical programs, tapes, and data were used in all cases, this one was very difficult to compute. Several times during the computations, extremely large relaxation residuals occurred. Also, on two separate occasions, the computer stopped suddenly after having completed many time steps successfully. The reason for this stop could not be determined, and it was assumed to be due to machine error. If so, it is possible that the final results were influenced by some unknown machine error and are not realistic. No unusual trouble was experienced in any of the other computations.

C. An additional pair of computations were made in which a strip of observed data was inserted across the 6×6 hole at each observation time. The first computation was made with the strip normal to the wind, the second with the strip along the wind. This single strip, consisting of 6 points, does not provide any vorticity values in the hole. Its main influence is in providing additional boundary values when the approximate vorticity field

is relaxed to obtain the streamfunction field in the hole. The results are shown in figure 11.

These "strip" computations were made with the same data that were used for the 6×6 case (with a 12 hour frequency of observation) which is also included in figure 11. Comparing these two results, we see that the introduction of this strip resulted in a reduction of error initially and at later times. The case where the strip was normal to the wind increased the error slightly between 12 and 24 hours but showed a consistent decrease of error elsewhere and the error was still decreasing at 48 hours. The case where the strip was along the wind, however, showed a consistent increase in error after 24 hours.

Both cases were definitely better than in the case of the hole without the strip and, unless the error growth continued in the second case, they were even better than the 3×3 cases. This is a little surprising because the hole with the strip still had 30 points where the data were unknown while the small hole had only 9. Also, the initial RMS wind error for the strip case was almost one m.p.s. larger than for the small hole. Apparently, for very small holes (of the order of 1000 km or less) the circulation and smoothing technique for obtaining the initial vorticity field results in a fairly close approximation to the true values and further improvement is difficult. We are not advocating the establishment of large holes but are merely pointing out that, where large natural holes already exist (oceans, uninhabited areas, etc.), the judicious application of this technique could be helpful. If additional observing sites were to be established, then these results (or other results based on tests in the area in question) could indicate the most advantageous sites to be chosen.

The results of these strip experiments suggest that aerial reconnaissance would be an excellent method for obtaining observations. In areas with sufficient air traffic, these observations could be made on a routine basis by commercial and other regular aircraft. In areas with little or no traffic, this function could be performed by special weather reconnaissance aircraft, just as is done at the present time. The latter have a distinct advantage in that the tracks flown can be varied on a seasonal (or synoptic situation) basis so as to produce the best results.

D. A statistical analysis of the operational

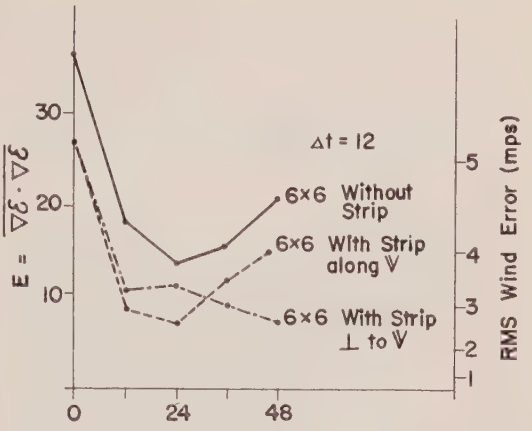


Fig. 11. Change of E with time for a 2100×2100 km hole with one row of observed data across the hole. The upper curve is for the same hole without the strip.

computation results is shown in Table 1. The actual mean wind velocity $\mathbf{V}'$ in a square hole will be larger than the mean normal velocity by an average factor of $\frac{\pi}{2}$. Based on this value and the values of errors in the predicted field in the hole, the root mean square error of wind velocity σ_v, and the root mean square error of wind direction σ_Θ (assuming no error in speed) were computed.

Table 1

Case*	$\mathbf{V}'$ (mps)	σ_v (mps)	σ_Θ (degrees)
$6 \times 6_{24}$	11.5	4.7	23.6
$6 \times 6_{12}$	11.5	4.0	20.0
$6 \times 6_{12}$ with strip $\perp$ to $\mathbf{V}$	11.5	3.1	15.5
$3 \times 3_{24}$	10.8	4.3	23.0
$3 \times 3_{12}$	10.8	3.5	18.7

* The case notation is the same as in figures 10 and 11.

4. Conclusions

It is evident that the method described in the paper does result in an overall decrease of error in region of low density observations. It is felt, therefore, that it should be applied by those organizations who are engaged in routine large scale weather analysis and forecasting by numerical methods. It is also evident that the 12 hour frequency of observation was

superior to the 24 hour frequency during these tests. It would have been interesting to make similar tests using observations at more frequent intervals to determine this effect more accurately. Unfortunately, however, such data over an adequately large area could not be located.

The largest hole used during these tests was approximately half as large as the natural hole in the North Atlantic Ocean. If we assume that similar results would be obtained in a hole as large as the North Atlantic, it appears that with sufficient boundary data surrounding this hole, we could expect an accuracy of the order of 5 m.p.s. RMS wind error. We were not able to carry out tests involving holes that large, partly because of the limited capacity of the BESK computer and partly because we could not locate a hole that large (in the data fields we used) with sufficiently dense and extensive known data on the boundaries.

It was noted that, generally, there was no great trend for improvement in the results after 24 hours. Assuming the actual advective wind velocity to be 10 m.p.s., advection from the inflow boundary would only penetrate about 800 km into the hole in 24 hours. This suggests that boundary influence and "region of influence" effects also contribute in reducing the error.

It is interesting that the oscillation of the error had a period of approximately 24 hours. VEDERMAN AND HUBERT (1957) found a similar period in the non-conservation of absolute vorticity as computed by a finite difference technique. There appears to be no physical explanation for this. Perhaps it is an inherent property of the finite difference method which has not yet been isolated.

6. Suggestions for Further Investigation

During the course of these experiments, a number of ideas, suggestions, and questions were brought up. Some were discarded, some were attempted, and some were reluctantly by-passed because they were beyond the capability of the system we used. They are listed here for those who plan further work in this subject.

1. Extend the tests to larger size holes.
2. Test the effects of asymmetric holes.
3. Try shorter observation frequencies (either by generating a hypothetical set of observations or by using data such as the "Tornado Alley" observations in the U.S.).
4. Do experiments similar to the strip cases with only single points or small triangles (from which the vorticity can be computed) of known data in the hole.
5. Require that the mean vorticity in the hole agrees with circulation around the hole at all times, not just initially.
6. Trace the advection during each time step, make a new circulation calculation around that part of the hole which has not been affected by the advection, and revise the approximate vorticity field accordingly.

Acknowledgement

I would like to express particular gratitude to Col. P. D. Thompson, who suggested these experiments originally and also provided inestimable guidance and enlightenment while they were being carried out. Special thanks are also due to Dr. F. B. Smith, for his contributions to the theoretical section. I also acknowledge my debt to Bengt Bohman, who collaborated in writing the numerical programs and operated the computer. I also thank all those others at the International Meteorological Institute who contributed by suggestions and discussions, especially Dr. Bert Bolin and Fil. lic. Ingemar Holmstrom.

Acknowledgement is also due to the Research and Development Command, United States Air Force, who supported most of the numerical computations and to the Royal Swedish Air Force, for supporting part of the computation.

REFERENCES

THOMPSON, P. D., 1961: A dynamical method of analyzing meteorological data *Tellus* **13**, pp. 334—349.

VEDERMAN, J. AND HUBERT, W. E., 1957: On the conservation of finite-difference absolute vorticity in barotropic forecasts. *Monthly Weather Review*, June 1957, pp. 193—199.

WELANDER, P., 1956: *Analysis of the 500-mb-surface over a region of no data* (unpublished). Report prepared for Royal Swedish Air Force, Stockholm.

On the Effect of Friction in Baroclinic Waves[1]

By E. O. HOLOPAINEN, International Meteorological Institute, Stockholm.

(Manuscript received July 10, 1961)

Abstract

Pertubation equations for a two-layer model are used to study the effect of the frictiona boundary layer in waves on a baroclinic current. It is found, that the change of the propagation speed of waves due to friction is significant only within the waveband, for which amplification occurs. Compared with the frictionless case, this waveband is broader when friction is included. The frictional instability is explained by kinetic energy considerations.

1. Introduction

The atmosphere does work against the surface of the earth and thereby loses kinetic energy. In addition internal friction transforms kinetic energy into internal energy.

This dissipative effect of friction together with heating are factors which determine the long-term behaviour of the atmosphere. However, for short period numerical prediction a more important effect of friction is probably that it generally causes convergence and divergence near the earth surface, where the friction force is comparable with the pressure and coriolis forces. Through the vertical velocities thus generated, the frictional influence can be transferred to those parts of the atmosphere where local friction forces are negligible.

CHARNEY and ELIASSEN (1949) have studied the effects of the inclusion of the frictionally created vertical velocities into an equivalent barotropic model. In such a model the only influence of friction is to dissipate kinetic energy. However, in a baroclinic atmosphere this need not be true. The frictionally-created vertical motions can be associated with a thermal field in such a way that kinetic energy is released. This is the case for example in a tropical cyclone.

In this note, perturbation equations for a two-layer model are used to study the effect of friction in waves on a baroclinic current. The intention is to get a qualitative picture of the effect of incorporating friction into this model in some reasonable way.

2. The Model

Two adjacent layers of the atmosphere are considered, lying between pressure levels p_0, p_2 and p_4,

where $$p_0 - p_2 = p_2 - p_4 > 0.$$

The mean pressure levels in these ayers are indicated by indices 1 and 3. The pressure surface p_0 is assumed to coincide approximately with the top of the frictional layer, and p_4 must be such that at that level $\omega = \frac{dp}{dt}$ can be assumed zero. The equations for this model are (e.g. THOMPSON 1961)

[1] The research reported here has been sponsered in part by the Geophysics Research Directorate, AFCRL, AFRD, of the Air Research and Development Command, USAF, through its European Office.

$$\frac{\partial}{\partial t}\nabla_p^2 z + \frac{g}{f}J(z,\nabla^2 z) + \beta\frac{\partial z}{\partial x} - \frac{f^2}{g}\frac{\partial \omega}{\partial p} = 0 \qquad (1a)$$

$$\frac{\partial}{\partial t}\left(\frac{\partial z}{\partial p}\right) + \frac{g}{f}J\left(z, \frac{\partial z}{\partial p}\right) - \sigma\omega = 0.$$
$$\sigma = \frac{\alpha}{g} - \frac{\partial \ln \Theta}{\partial p}; J(z,z') = \frac{\partial z}{\partial x}\frac{\partial z'}{\partial y} - \frac{\partial z}{\partial y}\frac{\partial z'}{\partial x} \qquad (1b)$$

As usual z is the height of the isobaric surface, g is the acceleration due to gravity, f is the coriolis parameter and $\beta = \frac{\partial f}{\partial y}$, which is assumed constant. Θ is the potential temperature and α the specific volume.

Eqs. (1a) and (1b) comprise a complete system of equations with two dependent variables z and ω.

3. Derivation of the phase-speed equation

The motion is assumed to consist of small perturbations superimposed on a basic current $U(p)$, a function of p alone. The perturbation field is assumed independent of y. Since it is implicitly assumed that the effect of dissipation on the basic current is exactly balanced by some external energy input, consideration of the effect of friction in our model may be confined to the perturbation field.

If z and ω now refer to the corresponding perturbation quantities, the following equations are obtained from (1):

$$\frac{\partial}{\partial t}\frac{\partial^2 z}{\partial x^2} + U\frac{\partial^3 z}{\partial x^3} + \beta\frac{\partial z}{\partial x} - \frac{f^2}{g}\frac{\partial \omega}{\partial p} = 0. \qquad (2a)$$

$$\frac{\partial}{\partial t}\frac{\partial z}{\partial p} + U\frac{\partial}{\partial x}\frac{\partial z}{\partial p} - \frac{dU}{dp}\frac{\partial z}{\partial x} - \sigma\omega = 0. \qquad (2b)$$

The vorticity equation (2a) is applied to the levels 1 and 3, the thermodynamic equation (2b) at the level 2. The vertical derivatives $\frac{\partial z}{\partial p}$ and $\frac{\partial \omega}{\partial p}$ are estimated by finite differences. The stability factor σ is taken to be constant. At the level p_0, ω is assumed to have some value ω_0 dependent on the mass convergence or divergence in the frictional layer.

The boundary conditions are thus:

$\omega = \omega_0$ at the level 1.
$\omega = 0$ at the level 4.

The finite difference approximations to the p-derivative of ω become:

$$\frac{\partial \omega_3}{\partial p} = \frac{\omega_2}{\Delta p}, \quad \frac{\partial \omega_1}{\partial p} = \frac{\omega_0 - \omega_2}{\Delta p},$$

where Δp is the pressure thickness of each of the two layers.

Instead of the heights z_1, z_3 and the velocities U_1, U_3 new variables are introduced, defined by:

$$\bar{z} = \frac{1}{2}(z_3 + z_1); \quad \bar{U} = \frac{1}{2}(U_3 + U_1)$$

$$z^\star = \frac{1}{2}(z_3 - z_1); \quad U^\star = \frac{1}{2}(U_3 - U_1).$$

$\bar{z}$ and $\bar{U}$ can be considered as the height and the basic current, respectively, at the pressure level p_2. To the same order of approximation, $z^\star$ is proportional to the temperature of the perturbation field and $U^\star$ to the vertical shear of the basic current, at this level.

With this notation the following equations are obtained:

$$\frac{\partial}{\partial t}\frac{\partial^2 \bar{z}}{\partial x^2} = \bar{U}\frac{\partial}{\partial x}\frac{\partial^2 \bar{z}}{\partial x^2} + U^\star\frac{\partial^2 z^\star}{\partial x^2} + \beta\frac{\partial \bar{z}}{\partial x} - \frac{1}{2}\frac{f^2}{g}\frac{\omega_0}{\Delta p} = 0 \qquad (3a)$$

$$\frac{\partial}{\partial t}\frac{\partial^2 z^\star}{\partial x^2} + \bar{U}\frac{\partial}{\partial x}\frac{\partial^2 z^\star}{\partial x^2} + U^\star\frac{\partial}{\partial x}\frac{\partial^2 \bar{z}}{\partial x^2} + \beta\frac{\partial z^\star}{\partial x} - \frac{f^2}{g}\frac{\omega_2}{\Delta p} + \frac{1}{2}\frac{f^2}{g}\frac{\omega_0}{\Delta p} = 0. \qquad (3b)$$

$$\frac{\partial z^\star}{\partial t} + \bar{U}\frac{\partial z^\star}{\partial x} - U^\star\frac{\partial \bar{z}}{\partial x} + \sigma\omega_2\frac{\Delta p}{2} = 0. \qquad (3c)$$

With the assumptions of the Ekman layer theory, the vertical velocity at the top of the friction layer has been found (e.g. CHARNEY and ELIASSEN, 1949) to be proportional to the geostrophic vorticity at that level:

$$w_0 = F\frac{\partial^2 z_0}{\partial x^2}, \text{ where } w = \frac{dz}{dt} \text{ and}$$

$$F = \frac{g}{f}\sin 2\gamma\sqrt{\frac{\bar{K}}{2f}}$$

γ is the angle between the surface wind and isobars, and K is the eddy viscosity.

Since the total rate of change of pressure arises mainly from vertical motion, we may write:

$$\omega_0 = -\varrho_0 g w_0 = -\varrho_0 g F \frac{\partial^2 z_0}{\partial x^2}.$$

The quantity $\frac{\partial^2 z_0}{\partial x^2}$ is obtained by linear extrapolation:

$$\frac{\partial^2 z_0}{\partial x^2} = \frac{\partial^2 \bar{z}}{\partial x^2} - 2\frac{\partial^2 z^\star}{\partial x^2},$$

When the expression for ω_0 is introduced into eqs. (3) and ω_2 eliminated between (3b) and (3c), the following equations result

$$\frac{\partial}{\partial t}\frac{\partial^2 \bar{z}}{\partial x^2} + \bar{U}\frac{\partial^3 \bar{z}}{\partial x^3} + U^\star \frac{\partial^3 z^\star}{\partial x^3} + \beta \frac{\partial \bar{z}}{\partial x} + \\ + \frac{k}{2}\frac{\partial^2 \bar{z}}{\partial x^2} - k\frac{\partial^2 z^\star}{\partial x^2} = 0 \qquad (4a)$$

$$\frac{\partial}{\partial t}\frac{\partial^2 z^\star}{\partial x^2} + \bar{U}\frac{\partial^3 z^\star}{\partial x^3} + U^\star \frac{\partial^3 \bar{z}}{\partial x^3} + \beta \frac{\partial z^\star}{\partial x} - \\ - \mu^2\left(\frac{\partial z^\star}{\partial t} + \bar{U}\frac{\partial z^\star}{\partial x} - U^\star \frac{\partial^2 \bar{z}}{\partial z}\right) - \\ - \frac{k}{2}\frac{\partial^2 \bar{z}}{\partial x^2} + k\frac{\partial^2 z^\star}{\partial x^2} = 0, \qquad (4b)$$

where $\mu^2 = -\frac{2f^2}{\sigma g(\Delta p)^2}$ and the constant k contains F.

These equations are homogeneous linear differential equations in $\bar{z}$ and $z^\star$. An elementary solution of the form

$$\bar{z} = \bar{M} \cdot e^{i\alpha(x-ct)}$$
$$z^\star = M^\star \cdot e^{i\alpha(x-ct)} \qquad \text{may be sought.}$$

α is a constant wave number and c is the phase-speed. Introducing these relationships into eqs. (4), we obtain:

$$\left\{i[\alpha^2(c-\bar{U}) + \beta] - \frac{k\alpha}{2}\right\}\bar{z} - \{i\alpha^2 U^\star - k\alpha\}z^\star = 0 \qquad (5a)$$

$$\left\{iU^\star(\mu^2 - \alpha^2) + \frac{k\alpha}{2}\right\}\bar{z} + \\ + \{i[(\alpha^2 + \mu^2)(c-\bar{U}) + \beta] - k\alpha\}z^\star = 0. \qquad (5b)$$

A non-zero solution for $\bar{z}$ and $z^\star$ exists only if the determinant of the coefficients vanishes. This condition gives the following quadratic equation for $c - \bar{U}$:

$$A_0\,(c-\bar{U})^2 + (A_1 + iB_1)(c-\bar{U}) + \\ + (A_2 + iB_2) = 0, \qquad (6)$$

where

$$A_0 = \alpha^2(\alpha^2 + \mu^2)$$
$$A_1 = \beta(2\alpha^2 + \mu^2)$$
$$A_2 = \frac{k\alpha}{2}(3\alpha^2 + \mu^2)$$
$$B_1 = \alpha^2 U^{\star 2}(\mu^2 - \alpha^2) + \beta^2$$
$$B_2 = \frac{k\alpha}{2}[U^\star(2\mu^2 - 3\alpha^2) + 3\beta]$$

The general solution to (4) can then be written as

$$\bar{z} = \sum_{\alpha=0}^{\infty}(\bar{M}_1 e^{ic_1\alpha t} + \bar{M}_2 e^{ic_2\alpha t})e^{i\alpha x} \qquad (7a)$$

$$z^\star = \sum_{\alpha=0}^{\infty}(M_1^\star e^{ic_1\alpha t} + M_2^\star e^{ic_2\alpha t})e^{i\alpha x} \qquad (7b)$$

Equation (6) may be written in a non-dimensional form in terms of $x = \frac{\alpha}{\mu}$, $S = \frac{U^\star}{\beta/\mu^2}$ and $\varkappa = \frac{k}{\beta/\mu}$. The folowing form results

$$\frac{C-\bar{U}}{\beta/\mu^2} = -a - ib \pm \sqrt{a^2 - d - b^2 + i(e-f)}, \qquad (8)$$

where

$$a = \frac{x^2}{2}\frac{2+x^2}{1+x^2}$$

$$b = \frac{\varkappa}{4}\frac{x(3+x^2)}{1+x^2}$$

$$d = \frac{x^4 + S^2(x^2-1)}{1+x^2}$$

$$e = \frac{\varkappa}{4}\frac{x^3(2+x^2)(3+x^2)}{(1+x^2)^2}$$

$$f = \frac{\varkappa}{2}x \cdot \frac{(2S+3)x^2 - 3S}{1+x^2}.$$

The real part of the phase velocity (8) is the speed of propagation of the perturbations, the imaginary part multiplied by the corresponding wave number is the growth rate of the wave.

4. The effect of friction

In Figure 1, the speed of propagation of waves, corresponding to $S=4$ and $\varkappa=1$, are given as a function of the wavelength X. This value of S has been chosen in order to demonstrate all the characteristic features of the frictionless case. This case has been discussed, for example, by Eliassen (1952).

The chosen value of $\varkappa$ corresponds to a typical eddy viscosity of the order of 10^5 cm^2 sec^{-1}.

If the imaginary part c_i of the phase speed is positive, the waves are growing. In this case, c_i is shown in Figure 1.

The characteristic features of the solution are:

a) *Frictionless case*

The waves consist of two components, which outside a specified band of wave numbers are neutral (constant amplitude) and have different speeds of propagation. Within the wavelength band, both components have the same velocity relative to the basic current, but one component is amplifying while the other is damped. This waveband is determined by the vertical shear of the basic current and the stability of the atmosphere.

It can be shown, that in the case of neutral waves, the geopotential and thermal fields are in phase or 180° out of phase, but for amplifying waves the thermal field lags behind the geopotential field. The physical explanation for such an amplification is that kinetic energy is produced by the solenoidal circulation set up by the associated thermal and ω-fields.

b) *Friction included*

The effect of friction on the propagation speed of waves is only markedly significant within the waveband, for which amplification occurs. For the longer amplifying waves in this band, friction results in propagation speeds closer to those for barotropic Rossby waves.

Compared with the frictionless case, the effect of friction reduces the amplitude of the perturbation field except within two narrow wavebands. These remarkable exceptions arise from a broadening of the waveband, for which amplification occurs. To wholly explain this

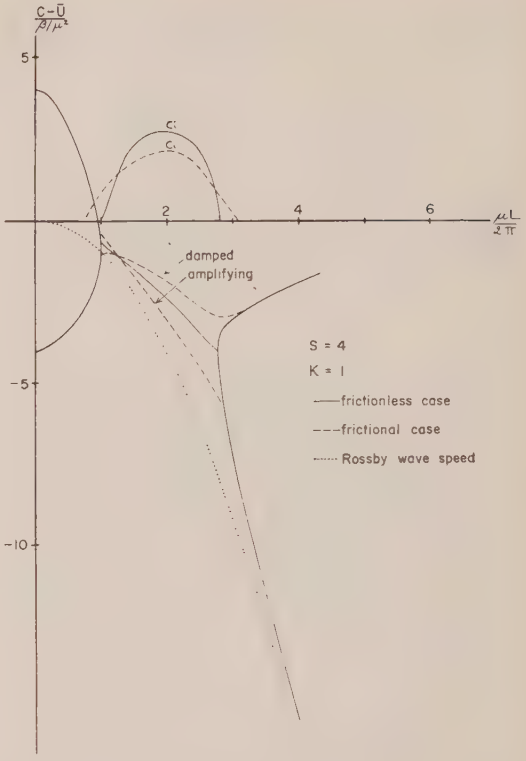


Fig. 1. The real part and the positive imaginary part of the phase speed $\frac{c-\bar{U}}{\beta/\mu^2}$ as a function of a dimensionless wavelength $\frac{\mu L}{2\pi}$.

phenomenon one has to study the three-dimensional structure of the waves described by the general solution (7) subject to certain initial conditions. However, a qualitative explanation can be obtained more easily by considering changes in the kinetic energy, as given by the model.

5. Kinetic energy considerations

From the physical point of view, the instability of perturbations may be regarded as a consequence of energy conversion in which internal and potential energy of the atmosphere is transformed into kinetic energy of the perturbations.

The perturbation kinetic energy in our model can be expressed as

$$K = \frac{1}{2}(v_3^2 + v_1^2) = \bar{v}^2 + v^{\star 2} =$$

$$= \frac{g^2}{f^2}\left[\left(\frac{\partial \bar{z}}{\partial x}\right)^2 + \left(\frac{\partial z^\star}{\partial x}\right)^2\right],$$

where v is the perturbation (meridional) wind and $\bar{v} = {}^1/_2\,(v_3 + v_1)$, $v^\star = {}^1/_2\,(v_3 - v_1)$.

An expression for the change of the average kinetic energy will be derived. Firstly:

$$\frac{\partial K}{\partial t} = \frac{2g^2}{f^2}\left[\frac{\partial \bar{z}}{\partial x}\frac{\partial}{\partial t}\left(\frac{\partial \bar{z}}{\partial x}\right) + \frac{\partial z^\star}{\partial x}\frac{\partial}{\partial t}\left(\frac{\partial z^\star}{\partial x}\right)\right].$$

Averaging over one whole wavelength yields

$$\frac{\widetilde{\partial K}}{\partial t} = -\frac{2g^2}{f^2}\left[\widetilde{z\frac{\partial}{\partial t}\frac{\partial^2 \bar{z}}{\partial x^2}} + \widetilde{z^\star \frac{\partial}{\partial t}\frac{\partial^2 z^\star}{\partial x^2}}\right],$$

where $\widetilde{(\quad)} = \frac{1}{L}\int_0^L (\quad)\,dx.$

Application of eqs. (3) finally gives

$$\frac{\widetilde{\partial K}}{\partial t} = -4\frac{g^2}{f^2}U^\star \frac{\alpha^2\mu^2}{\alpha^2+\mu^2}\widetilde{\bar{z}\frac{\partial z^\star}{\partial x}} - \frac{g^2}{f^2}\alpha^2 k \cdot \widetilde{\bar{z}(\bar{z} - 2z^\star)}. \qquad (9)$$

The first term on the right hand side indicates the general condition for the rate of change of the average kinetic energy of the perturbations in a frictionless case. The term is positive when $z^\star$-field lags behind the z-field, and has a maximum when this lag is 90° (THOMPSON, 1959).

Considering the second term, the vorticity at the pressure surface p_0 was obtained by a linear extrapolation

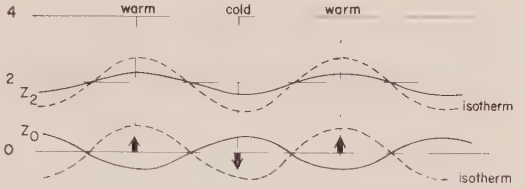


Fig. 2. Maximal growth of the kinetic energy due to the friction.

$$\frac{\partial^2 z_0}{\partial x^2} = \frac{\partial^2 \bar{z}}{\partial x^2} - 2\frac{\partial^2 z^\star}{\partial x^2}.$$

Consequently the z_0 and $(\bar{z} - 2z^\star)$ fields must be in phase. From (9) it follows that friction causes a decrease in the kinetic energy in the model, if the correlation between $\bar{z}$ and z_0 is positive, and a growth of the kinetic energy, if it is negative.

When using linear extrapolation, the quantity $\frac{\partial}{\partial p}\frac{\partial^2 z}{\partial x^2} = \frac{\partial^2}{\partial x^2}\frac{\partial z}{\partial p} = -\frac{R}{pg}\frac{\partial^2 T}{\partial x^2}$ is assumed to be constant in the vertical. In other words the temperature field is assumed to have no tilt in the vertical.

The condition for maximum growth of kinetic energy due to friction is schematically shown in Figure 2. The phase difference between $\bar{z}$-field and z_0-field is 180°. The arrows indicate the direction of the vertical movement at the level p_0.

The figure explains qualitatively what structure of waves is required for frictional instability. Because the $\bar{z}$ and $z^\star$-fields are in-phase, the first term on the right hand side of equation (9) is zero. The frictionally created vertical motions cause an extra positive solenoidal circulation and the mean kinetic energy can grow.

REFERENCES

BRUNT, D., 1941: *Physical and Dynamical Meteorology*. University Press, Cambridge.

CHARNEY, J. G. and ELIASSEN, A., 1949: A Numerical Method for Predicting the Perturbations of the Middle Latitude Westerlies. *Tellus* **1**, pp. 38—54.

ELIASSEN, A., 1952: Simplified Models of the Atmosphere, Designed for the Purpose of Numerical Weather Prediction. *Tellus* **4**, pp. 145—156.

THOMPSON, P. D., 1959: Some Statistical Aspects of the Dynamical Processes of Growth and Occlusion in Simple Baroclinic Models. *The Rossby Memorial Volume*, New York.

THOMPSON, P. D., 1961: *Numerical Weather Analysis and Prediction*. McGraw-Hill Book Company.

A Boundary Layer Interpretation of the Low-level Jet

By H. WEXLER, U.S. Weather Bureau, Washington, D.C.

(Manuscript received June 27, 1961)

Abstract

The inertial boundary layer theory developed by Charney and Morgan to explain the intensification of currents near the western boundary of an ocean is applied to the northward flowing atmospheric jet stream observed in the lowest two kilometers over the central United States; here the mountainous spine of Central America serves as a barrier to the westward flowing trade-wind current and deflects it to the north. Because of diurnal frictional changes caused by surface heating and cooling, the jet undergoes a marked diurnal variation with a noctural maximum so strong that the current becomes super-geostrophic. The dynamic consequences of the unbalanced current are examined in the light of theories by Tepper and Veronis.

The thin, narrow concentration of momentum in the lowest 2 kilometers of the atmosphere known as the "low-level jet stream" has attracted surprisingly little attention, apart from work by WAGNER (1939), MEANS (1952, 1954), GIFFORD (1953), BLACKADAR et al. (1957), ESTOQUE (1959). This jet, only a few hundred kilometers wide, may move twice as fast as layers several hundred meters below or above (figs. 1, 11). Although its direction is fairly constant from the south or southwest, its speed has a large diurnal variation with an early morning maximum double or more the afternoon minimum (fig. 2). In the United

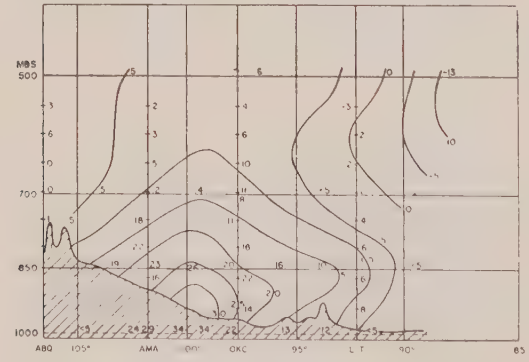


Fig. 1. Mean southerly wind components along 35 degrees latitude, 1500Z July 10th through 0300Z July12 th, 1951, speed in knots (after Means, 1954),

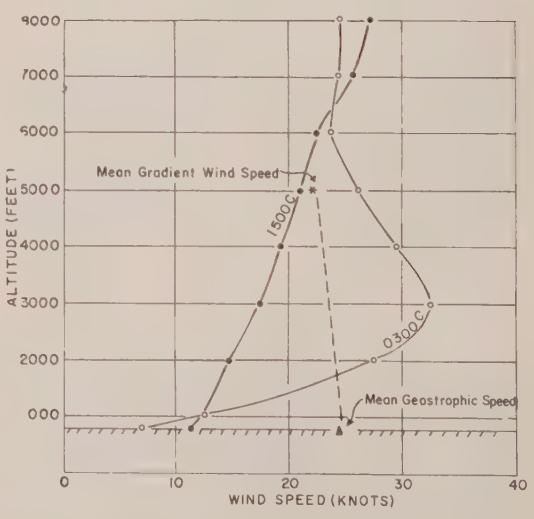


Fig. 2. Average wind speed profile for 16 significant boundary layer jets at San Antonio, Texas, during January 1953 (after Blackadar et al., 1957).

States, it appears most prevalent in the strip between 95° W and 100° W, from southern Texas to Nebraska, a region also noted for its high frequency of severe nocturnal thunder- and wind-storms. Pilots are grateful for its 30 knots or higher "sunrise" tail-winds when flying north and avoid its opposing winds on return flights (POLSON, 1958). On the other hand, the strong vertical wind shears accompanying the jet are often dangerous to aircraft descending for nocturnal landings at near-stall speeds; practical procedures for forecasting this phenomenon have been devised by BLACKADAR and REITER (1958).

WAGNER (1939) attempted to explain this nocturnal acceleration of the low-level winds by drainage of cold air from the high ground to the west, as a result of radiational cooling. NEWTON (1959) has suggested that the low-level jet is quite similar to the Gulf Stream in its narrow concentration of momentum and location just to the east of a meridional barrier —the slopes leading up to the Rocky Mountains in the case of the atmospheric jet and the continental shelf for the Gulf Stream.

In this paper we shall examine further this analogy by applying to the atmospheric jet the inertial boundary layer theories developed and applied so successfully to the Gulf Stream by CHARNEY (1955) and MORGAN (1956) and shall utilize the unbalanced current models of TEPPER (1955) and VERONIS (1956) to study the dynamic consequences of the observed large diurnal variation of the jet from sub- to super-geostrophic speeds.

Western Boundary Currents of Oceans

The intensification of currents near western boundaries of oceans such as the Gulf Stream and the Kuroshio has been the subject of intensive research in recent years. A good review is given by STOMMEL (1958), to whom credit is given for first pointing out the critical role played by the latitudinal variation of the Coriolis parameter in the formation of such jet-like currents (1948). The linear viscous theories of STOMMEL (1948), MUNK (1950), HIDAKA (1949) explained the principal features of the mean circulation but left unexplained the details of the individual boundary layer. This was examined by CHARNEY (1955) and MORGAN (1956) who retained in a narrow layer next to the western wall the important non-linear inertial terms neglected in the earlier models, but omitted the frictional terms which were presumed to be important only in the immediate vicinity of the wall.

A simple qualitative description of their results based on absolute vorticity conservation has often been used to explain variations in shear and curvature of currents moving in north-south directions. If $\frac{f+\zeta}{D}$ is constant for individual columns moving northward, then as f, the Coriolis parameter, increases with increasing latitude, ζ, the relative vorticity and/or D, the column thickness, must decrease. If D is kept constant, the column must acquire more anticyclone vorticity relative to the earth. If this is converted mainly into anticyclonic shear, then there must develop a high-speed current at the western boundary of the flow.

Both Charney and Morgan divide the oceanic flow into two regions—a broad exterior current flowing westward and a narrow boundary layer adjacent to the vertical western wall. In the exterior region the flow is geostrophic but in the boundary layer where only pressure and inertial forces are present, the meridional component, v, of flow is geostrophic but not the zonal component, u, because of presence of the large terms, $u\frac{\partial v}{\partial x}$ and $v\frac{\partial v}{\partial y}$, where x points to the east and y to the north. By transforming the primitive equations into those expressing the conservation of potential vorticity and the Bernouilli equation, Charney and Morgan found the volume transport streamlines and surface contours both for the case of a single homogeneous layer and a double layer, where zero horizontal pressure gradient was assumed in the lower thicker layer. Charney assumed that in the interior region, the volume transport of the westward wind-driven surface layer decreased linearly to the north over a distance of 1,400 km, from the Florida Straits to Cape Hatteras; however, Morgan assumed a volume transport which did not vary northward in 2,000 km. Despite this difference in the meridional distribution of the westward transport of water and the difference in magnitude of transport, $8 \cdot 10^{13}$ cm^3 sec^{-1} assumed by

Charney and $2 \cdot 10^{13}$ cm^3 sec^{-1} by Morgan, their results are quite similar.

Application to the Atmosphere

In applying to the atmosphere the inertial boundary layer theory derived for the ocean, the continental slope off the east coast of the United States is replaced as the western wall by the high escarpment of Central America with an average height of about two kilometers. The wind-driven surface water transport is replaced by the trade-wind transport itself as shown schematically in figure 3. The trade-winds entering the western Caribbean Sea and the Gulf of Mexico impinge on the southeast-northwest oriented Isthmus of Panama; the northern portion is deflected northwestward and northward and takes on the characteristics of an inertial boundary layer because of the northward increase of the Coriolis parameter, and enters the Mississippi Valley. In the computations below, it is assumed that that portion of the trade-winds south of Jamaica crosses the Isthmus and does not enter the Mississippi Valley.

In Table 1 are shown the average monthly sea-level geostrophic wind speeds perpendicular to the 80° W meridian and also the average annual values for 1941—50 for the Colon—Key West section, the Kingston, Jamaica—Key West section, and the Kingston-Jacksonville section, computed from pressure

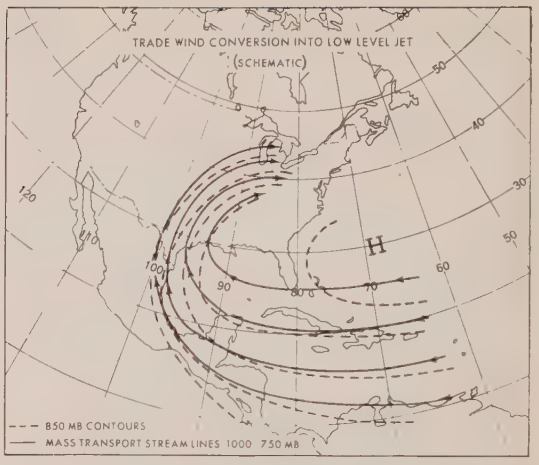


Fig. 3. Schematic sketch of trade-winds entering the Caribbean Sea, Gulf of Mexico and southern United States.

data published in World Weather Records, 1941—50 (Anonymous, 1959).

The average monthly geostrophic wind speeds are about the same across the line from Colon to Key West, and from Kingston to Key West, but smaller from Kingston to Jacksonville. During the Midwest "severe-storm" season, March to May, there is little variation in wind speed, and surprisingly little year-to-year variation in the average annual wind speeds.

Morgan's analysis is applied to a 2-layer atmosphere for two different widths of that

Table 1

Average sea-level geostrophic trade-wind speeds (meter sec^{-1}) across the 80th West Meridian from Colon (9°21′ N) to Key West (24°35′ N); from Kingston (18°15′ N) to Key West; and from Kingston to Jacksonville (30°20′ N).

	Average 1941—50			Year			
	CO-KW	*KI-KW*	*KI-JAX*		*CO-KW*	*KI-KW*	*KI-JAX*
Jan.	8.7	9.3	6.3	1941	7.0	7.3	4.6
Feb.	7.9	8.2	5.1	1942	6.3	7.7	5.1
Mar.	6.9	7.3	4.0	1943	6.8	7.3	4.7
Apr.	6.1	6.6	4.6	1944	6.8	7.3	4.7
May	6.0	6.1	3.8	1945	6.4	6.4	4.3
June	6.3	5.2	2.7	1946	6.8	6.6	4.8
July	7.0	5.4	2.9	1947	6.8	6.6	5.1
Aug.	5.6	5.2	3.2	1948	6.4	6.8	4.3
Sept.	3.8	3.6	4.4	1949	6.5	7.3	4.8
Oct.	4.2	4.3	6.1	1950	6.9	8.2	5.5
Nov.	7.1	9.8	6.9	Annual	6.5	6.8	4.7
Dec.	9.2	10.2	6.8				
Annual	6.5	6.8	4.7				

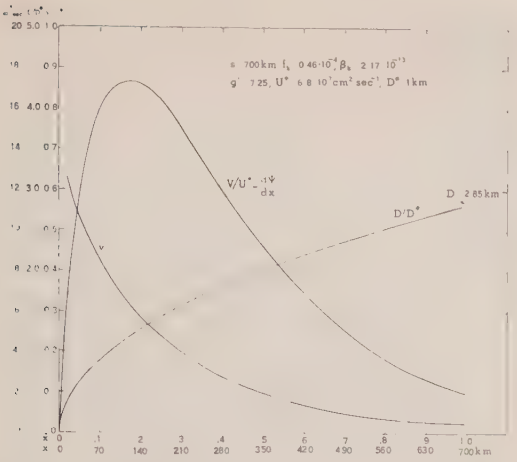


Fig. 4. Computed northward volume transport, (V), relative to trade-wind transport ($U^\star$), velocity (v) and ratio of depth of lower layer (D) to its initial depth ($D^\star$), at the northern boundary of the region where the theory applies, for Case I.

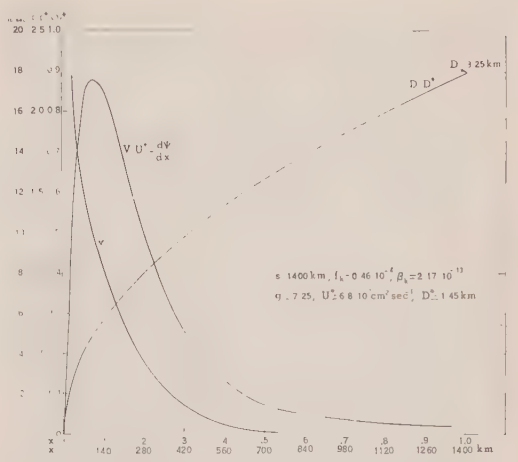


Fig. 5. Computed northward volume transport, (V), relative to trade-wind transport ($U^\star$), velocity (v) and ratio of depth of lower layer to its initial depth ($D^\star$), at the northern boundary of the region where the theory applies, for Case II.

portion of the trade-wind current (assumed to be north of Jamaica) which turns northward into the Mississippi Valley: Kingston, Jamaica to Key West, Florida (700 kilometers) and Kingston to Jacksonville, Florida (1,400 kilometers). In each case the westward transport of trade-wind air across a 1 cm wide strip is taken as $6.8 \cdot 10^7$ cm^2 sec^{-1}, a value corresponding to (I) an average annual sea-level geostrophic velocity perpendicular to a line from Kingston to Key West of 6.8 m sec^{-1} (see Table 1) and an assumed average depth of the trade-wind layer of 1 km, and (II) an average annual geostrophic velocity of 4.7 m sec^{-1} from Kingston to Jacksonville and an assumed depth of the trade-wind layer of 1.45 km. The total transport thus differs by a factor of 2.

When the deeper upper layer is assumed to be motionless, the same equations used for the Gulf Stream apply. The results for the two different transports are shown in figure 4 (Case I) and figure 5 (Case II) for the special condition of the lower layer vanishing at the northwestern corner of the region of applicability of the boundary layer theory.

For this special case, in Morgan's notation:

$$\varepsilon = \frac{g' D^{\star 2}}{\beta_k s^2 U^\star} = 1$$

where $\beta_k = 2.17 \cdot 10^{-13}$ cm^{-1} sec^{-1}, the gradient of the Coriolis parameter at the latitude of Kingston, 18°15′ N,

$U^\star = 6.8 \cdot 10^7$ cm^2 sec^{-1} the average westward volume transport of trade-wind air across a strip 1 cm wide,

s = 700 km or 1,400 km, and

$D^\star$ = 1 km or 1.45 km;

whence g' = 7.25 cm sec^{-2} or 14.50 cm sec^{-2}, corresponding to a temperature inversion of 2.3° C or 4.6° C, respectively, between the two layers.

In figure 4 are plotted three curves against the abscissa $\bar{x} = x/s$, which apply to the northern boundary of the flow ($\bar{y} = y/s = 1$) for the smaller trade-wind transport (Case I):

i) The northward volume transport, V, relative to the westward trade-wind transport $U^\star$. This transport, zero at the wall, reaches a maximum at 118 km east of the wall and approaches zero asymptotically as x increases.

ii) The absolute value of the northward velocity, v. This approaches infinity at $x = 0$, declines rapidly and asymptotically eastward, becoming less than 2 meters sec^{-1} at 350 km. The anticyclonic shears are quite

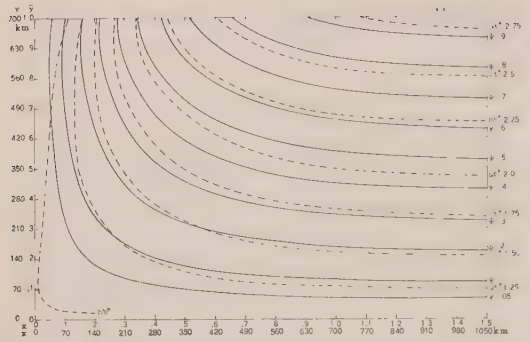


Fig. 6. Computed volume transport streamlines (full) and lines of equal D/D^* (dashed), Case I.

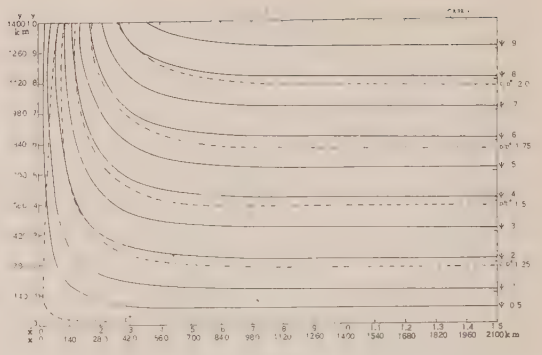


Fig. 7. Computed volume transport streamlines (full) and lines of equal D/D^* (dashed), Case II.

large—greater than or equal to the Coriolis parameter, f, for a distance 100 km east of the wall.

iii) The ratio of the depth of the lower layer, D, to the undisturbed depth of the trade-wind layer D^*. This increases rapidly from 0 at the wall to 2.85 at 700 km from the wall, where $D=2.85$ km.

In figure 5 the corresponding curves are applied to a trade-wind belt twice as wide (Case II), $s=1{,}400$ km, but having the same average unit volume transport as in the previous case, $U^*=6.8 \cdot 10^7$ cm^2 sec^{-1}, and a larger inversion of 4.6° C.

For the case of the wider trade-wind belt the maximum meridional transport in the boundary layer is 2.6 times larger than in the narrower case; it reaches a maximum at 105 km east of the wall, and is somewhat wider. The northward velocity is also larger than for the narrower trade-wind case, becoming less than 2 m sec^{-1} at 600 km from the wall. The anticyclonic shears are greater than or equal to f for a distance of about 150 km east of the wall. The D/D^* curve increases from zero at the wall to 2.25 at 1,400 km east of the wall where $D=3.25$ km.

In figures 6 and 7 are shown the volume transport streamlines (full) and lines of equal D/D^* (dashed) for the 700 km wide (Case I) and 1,400 km wide (Case II) trade-wind belts. The streamlines and contour lines are parallel and the motion is geostrophic beyond about 500 km east of the wall (outside of the inertial boundary layer). As the exterior current approaches the wall ($x=0$) from the right and is deflected northward, it turns not as a uniform current but as a strongly anticyclonic shearing current, increasingly so with distance westward to and northward along the wall. The volume transport streamlines cut across the contour lines towards lower height values and both sets of lines turn anticyclonically so that near the northern limit of domain of validity of the solution, they are directed slightly towards the east.

Computed and Observed Frictional Effects

It thus appears from the close agreement of observation and theory that the northward flow of air in the Great Plains can be interpreted as an inertial boundary layer similar to the Gulf Stream, characterized by strong lateral anticyclonic shears on the western flank. If lateral friction with the western wall is taken into account, as Neumann has done in a linear analysis of the Gulf Stream (1960), then there would be strong shears on the western side of the boundary flow so that the flow would have the appearance of a true jet in the horizontal, as indicated schematically by the dotted lines joining the "v" curves in figures 4 and 5. Vertical frictional stresses originating at the ground would create a vertical wind shear in the surface layer; above the surface layer, where the boundary layer acceleration does not occur, the speed is zero so that the flow also has the appearance of a jet in the vertical. As radiation operates alternately to create and destroy the surface nocturnal inversion, the vertical frictional stresses would undergo a strong diurnal varia-

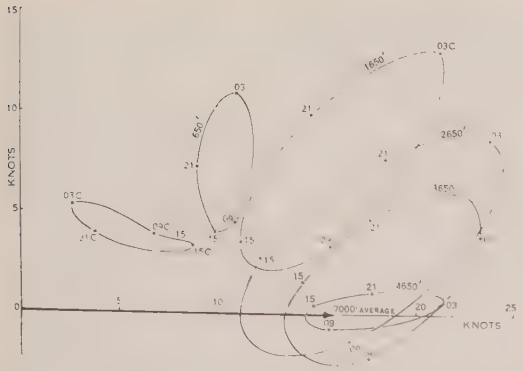


Fig. 8. Composite diagram of diurnal wind variation at Wichita, Kansas and Oklahoma City, Oklahoma (after Blackadar et al., 1957).

tion and so change greatly the profile of the vertical jet.

Blackadar et al. (1957) have constructed several mathematical models based on eddy coefficients varying with time of the day and height above the ground in attempts to reproduce the marked observed diurnal variation in the vertical jet. Despite ingenious combinations of time and height variations, the authors admit their inability to reproduce the magnitude of the diurnal change. It is believed that this deficiency in large measure is due to neglect of the mechanism which in the first place actually causes formation of the low-level jet itself. It is not the radiative and frictional effects which occur *locally* that create the highly concentrated flow in the western boundary layer, but the *bulk* properties of the flow caused by large scale inertial effects. This is not to deny the strong influence of large pressure systems moving in and out of the Midwest which might disrupt completely the flow of air from the south. But when southerly flow is present, as in the case when the western end of the Bermuda anticyclone covers the Mississippi Valley, then the inertial boundary layer influences enter in an important manner. Without this westward intensification of the southerly flow up the Mississippi Valley, there would not be a "basic flow" on which the frictional stresses could operate diurnally and so give the strong diurnal low-level jet so characteristic of that region.

During the warmest part of the day when frictional coupling of the low-level jet with the ground is greatest, the jet speed is considerably reduced from its value called for by the frictionless boundary layer theory. But at night, as the growing surface inversion increasingly shields the jet from surface friction, the jet speeds tend to be restored to their boundary layer values. During this interval, the wind speeds in the jet become super-geostrophic. This diurnal variation is illustrated in figure 8, taken from Blackadar et al. (1957), who made a composite diagram showing the average of the diurnal wind variations at different levels above the ground at Wichita, Kansas and Oklahoma City. The winds were taken from 6-hourly pilot balloon observations made during 29 days during the summer of 1951, selected to avoid marked changes of surface pressure gradient. The orientation of each wind was measured relative to the wind direction at 7,000 ft above sea-level, at which height the diurnal variation vanished. The diurnal variations are small at and near the ground, increase to a maximum at about 2,000 ft above the ground, at which level the wind speed is sub-geostrophic from about 0900 to about 1800 hours, local time, and super-geostrophic from 1800 to 0900 hours, with the maximum at about 0300 hours. In the sub-geostrophic phase, the wind velocity is such that there is balance between pressure gradient, Coriolis and frictional forces originating at the ground. In the nocturnal super-geostrophic phase, the winds are not in balance since the frictional forces are greatly diminished. Hence, during this phase the vector difference between actual and geostrophic wind describes a portion of an inertial oscillation of period 12 pendulum hours, which at latitude 37.5° is 19.7 hours. This oscillation moves the wind vector arrowheads in an anticyclonic path, theoretically a circle, but actually observed to be an ellipse, probably because of the presence of sizeable frictional forces during the warmer part of the day. The unbalanced flow which exists during the night hours, when the winds are released from the braking influence of ground friction, occurs over the entire jet-stream. The dynamic effects cannot be examined, therefore, merely as a problem involving only the wind changes over a single station, but must be studied from the view-point of the entire jet-stream.

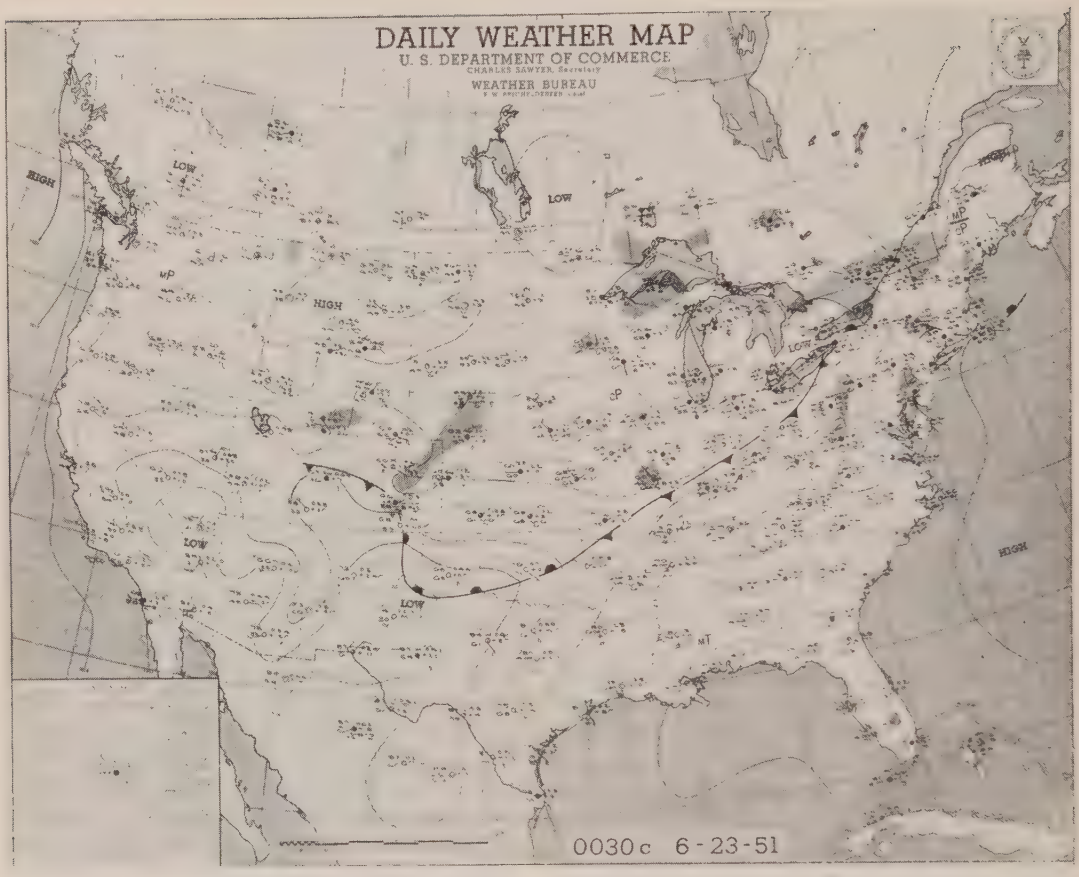


Fig. 9. Sea-level weather map for 0030 C.S.T., June 23, 1951.

Low-level Jet of June 23—24, 1951

A good example of the diurnal variation of the low-level jet occurred on June 23—24, 1951. The synoptic charts for 0030 C.S.T. for June 23 and 24 (figures 9 and 10) show the geostrophic flow of trade air around the western end of the Bermuda anticyclone which extends as far west as west Texas and as far south as latitude 10° N. The isobars indicate a stronger northward flow over central Texas than to the east and the air impinges on a stationary front extending from the Texas Panhandle east-northeast to northern Ohio. About 300 km south of this, five vertical cross-sections are drawn along the line from Albuquerque, New Mexico to Lake Charles, Louisiana (fig. 11). These sections, which reveal a low-level jet centered near Abilene, Texas, are approximately perpendicular to the jet axis and are drawn for 03, 09, 15, 21 and 03 hours C.S.T. The diurnal variation in the thermal stability of the surface layer is shown by the changes in the potential temperature isotherms (full or heavy dashed lines), while the strong variations in jet speed are shown by the isotachs (thin dashed lines). At 03 h when the stability is largest, the main jet axis is over Abilene, Texas, maximum speed 55 knots; at 09 h when the surface layer becomes unstable, this decreases to 35 knots; at 15 h, at maximum instability, the jet maximum decreases to 25 knots; at 21 h when the surface layer becomes more stable the jet maximum increases to 35 knots; and at 03 h the next morning the jet maximum increases to the 55 knot value observed 24 hours earlier. Presumably a major portion of the jet-stream becomes super-geostrophic during the night hours when frictional contact with the ground is curtailed.

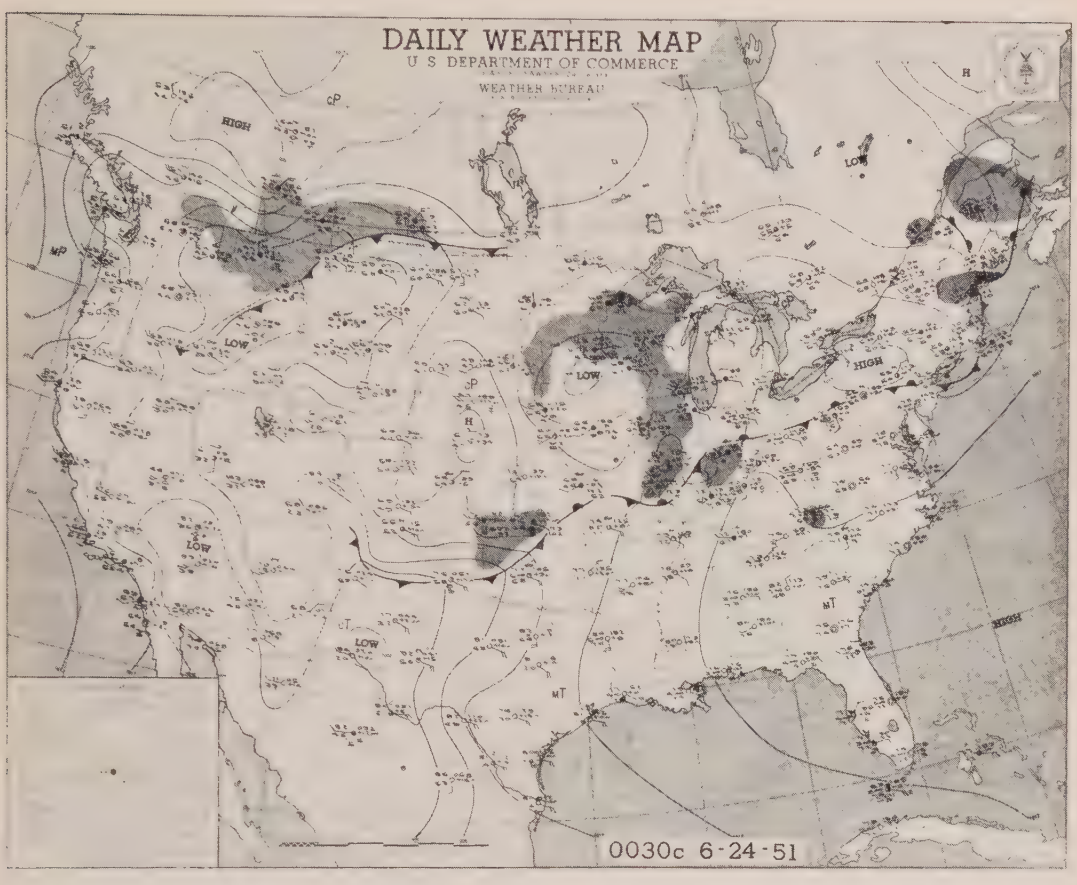


Fig. 10. Sea-level weather map for 0030 C.S.T., June 24, 1951.

There is fairly good agreement between some of the observed features of the low-level jet, illustrated in figure 11, and the theoretical features shown in figures 4 and 5.

First, we can define the top of the lower layer as the 305° Θ-isotherm, which is seen to intersect the surface just west of Big Spring, Texas in the bottom cross-section of figure 11 and rises eastward to a height of 1.7 km over Shreveport, Louisiana, 650 km to the east. This is smaller than 2.75 km at the same distance in Case I (figure 4) and close to the 1.5 km in Case II (figure 5).

Second, the observed jet-stream axis (center of the isotach maximum in the lower section of figure 11) is about 190 km east of the 305° Θ-isotherm intersection with the surface. If we define the computed jet-axis as the volume transport maximum, this is located 110 to 140 km east of the wall in figure 4 (Case I) and about 105 km east of the wall in figure 5 (Case II).

Third, the "width" of the observed jet, defined as the distance from the isotach maximum to the 5 knot (or 2.5 m sec^{-1}) isotach, is about 480 km in figure 11, bottom section. If we define the computed jet-width as the distance from the wall ($v = \infty$) to $v =$ 2.5 m sec^{-1}, the width is 300 km for Case I (figure 4) and 350 km for Case II (figure 5).

Dynamic Consequences of Unbalanced Flow in the Jet

We are thus lead to consider the dynamic consequences of an unbalanced jet-stream directed to the north. In 1955, at the writer's suggestion, such a problem was solved numerically for both a linear and non-linear model by Tepper (1955). Tepper assumed a

Gaussian-type profile jet-stream of width about 500 km and maximum speed of about 20 meters sec^{-1} in a 2.5 km thick surface layer separated by a 5° C temperature discontinuity from an upper layer where horizontal pressure gradients were assumed to vanish. With a value of $g' = g\frac{\varrho' - \varrho}{\varrho} = 17$, Tepper and Newstein also obtained numerical solutions for Gaussian-type profiles skewed to the east and west (1956).

A marked disturbance of the interface occurred as the Coriolis force caused the unbalanced current to move eastward in the initial phase of an inertial oscillation. The interface rose to the east of the jet axis and lowered to the west, with a maximum displacement upward at the end of 6 hours of about one kilometer at a distance of 500 km from the original jet axis. Results of both the linear and non-linear cases were in close agreement but with important exceptions: for the non-linear case the slope of the leading edge of the interface wave at its maximum amplitude was 1 : 275 compared to 1 : 450 for the linear case. Even more important, in the non-linear case, the slope steepened with time and became essentially vertical after 8 hours giving rise to a "pressure-jump", the atmospheric analog to the well-known hydraulic jump or shock wave.

In comprehensive meso-scale synoptic analyses, Tepper et al. (1954) showed that in northern Texas and southwestern Oklahoma, pressure jumps of lengths a few hundred kilometers and average duration of 3 to 4 hours, moved perpendicularly away from the basic current and to the right, at average speeds of about 20 meters sec^{-1}, a value close to $\sqrt{g'D}$, the theoretical propagation speed of gravity waves of wave-length large compared to the depth D. They were also found to be three times more frequent during the night hours (from 6 pm to 6 am local time) than during the daytime. This latter point is particularly important since, as noted earlier, the low-level jet stream is most strongly developed at night and in fact becomes super-geostrophic.

During the 8 month period from January to August 1951, 247 pressure jump lines swept through the dense network of 134 observing stations in southeastern Kansas and northern

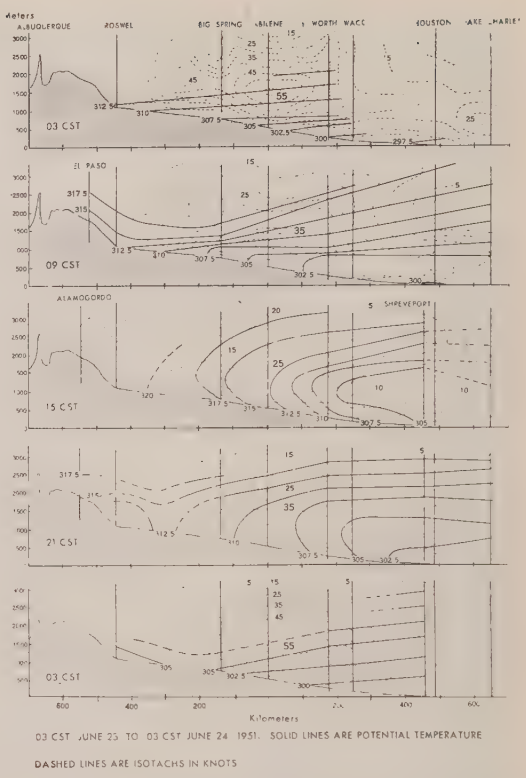


Fig. 11. Five west-east vertical cross-sections at 6-hourly intervals through the low-level jet stream showing the strong diurnal variation of thermal stability in the surface layer and in the speed of the jet, June 23—24, 1951.

Oklahoma and 725 or 86 % of all severe local storms reported were located in the areas swept out by some pressure jump line. It would appear that strong vertical stretching of the lower layer of air accompanied passage of the interface wave and with the air sufficiently humid, as was usually the case, there resulted condensation and liberation of latent heat of condensation which furnished much of the energy of the storm.

The presence of gravity waves on the interface and the steepening tendency of its leading edge are difficult to observe directly since a much denser network of aerological stations is required; but their existence is surmised from surface reports. However, the chain of evidence points to a definite connection between these nocturnal events: the formation of the surface inversion, the quenching of surface friction, the re-establishment of the strong

lew-lovel jet and the high frequency of line phenomena (pressure jumps and severe storms) propagating to the right (east or southeast) of the jet.

VERONIS (1956) pointed out two objections to Tepper's analysis of the unbalanced southerly jet: the assumption of an initially unbalanced current as large as 20 meters per second and the neglect of possible non-linear interactions between the baroclinic (or internal) oscillation mode which Tepper examined and the barotropic (or free surface) mode which he neglected. Veronis attempted to remove the first of these weaknesses by allowing the momentum to be given to the current over time intervals of 6 and 12 pendulum hours, respectively, but only analyzed both modes in a linear model so that interactions between internal and external modes were neglected.

Westerly momentum of amount $6\pi \cdot 10^4$ gm cm^{-1} sec^{-1} per unit volume was fed into a surface layer of ocean 200 km wide and the resulting axial and transverse motions, and perturbations of the free surface and the interface computed for $g' = 2$, a 500 meter thick top layer and a motionless bottom layer 3,500 meters thick. We may apply the numerical results of this model to the atmosphere where the 500 meter bottom layer is supplied with the momentum and the upper layer is motionless. The value of $g' = 2$ corresponds to a temperature inversion of about 0.6° C. Since the unit momentum given to the system occurs as a factor in each of the solutions, we can increase the oceanographic results to correspond with meteorological conditions. To do this, we decrease the density by a factor of 10^3 and thus increase the resulting wind speeds and interface and free surface displacements by the same factor. Also, we assume the model is applicable to a southerly current where variation of the Coriolis force with latitude is neglected; according to Veronis, this appears to be acceptable for the short time scales considered here.

Interpreting Veronis' results for a 500 meter thick lower atmospheric layer to which a momentum of amount $2.7 \cdot 10^7$ gm cm^{-1} sec^{-1} per unit volume is given over a period of 6 pendulum hours (nearly 10 hours at latitude 37.5°), we find that there are still evident rather violent lateral oscillations of 10 to 20 m sec^{-1} at the current center, with periods of about 12 pendulum hours. The surface pressure changes at the right edge of the current are of the order of 0.1 mb, or much smaller than observed pressure jumps. These changes are not as large as those found by Tepper who added momentum greater than the amount assumed here and did so instantaneously instead of over an interval of 6 pendulum hours. Also Tepper did not allow free surface oscillations to occur which would have compensated to some extent the pressure changes caused by variations of the interface.

Thus, the question as to whether some pressure jump lines derive their energy from the accelerating low-level jet as ground friction diminishes nocturnally cannot be answered now.

Acknowledgments

The writer acknowledges with gratitude stimulating discussions with Henry Stommel and George Veronis at the Woods Hole Oceanographic Institution and also the assistance provided by Walter Hoecker and Thomas H. Carpenter of the U.S. Weather Bureau in preparing figures and performing computations.

REFERENCES

ANONYMOUS, 1959: *World Weather Records, 1941—50*, U.S. Weather Bureau, Washington 25, D.C., 1361 pages.

BLACKADAR, A. K. et al., 1957: *Studies of wind structure in the lower atmosphere*, Dept. of Meteor. and Ocean., New York Univ., 164 pages.

BLACKADAR, A. K. and REITER, G. C., 1958: *Objective forecasting of low-level wind shear*, Scient. Rpt. No. 1, Dept. of Meteor., Penn. State Univ., 28 pages.

CHARNEY, J. G., 1955: The Gulf Stream as an inertial boundary layer, *Proc. Nat. Acad. Sci.*, **41**, pp. 731—740.

ESTOQUE, M. A., 1959: *A preliminary report on a boundary layer numerical experiment*, GRD Res. Note No. 20, USAF, Bedford, Mass.

GIFFORD, F., JR., 1952: The breakdown of a low-level inversion, etc., *Bull. Amer. Met. Soc.*, **33**, pp. 373—379.

HIDAKA, K., 1949: Mass transport in ocean currents and lateral mixing, *Geophys. Notes, Univ. of Tokyo*, **2**, pp. 1—4.

MEANS, L. L., 1952: On thunderstorm forecasting in the central United States, *Mo. Wea. Rev.*, **80**, pp. 165—189.

MEANS, L. L., — 1954: A study of the mean southerly wind-maximum in low levels associated with a period of summer precipitation in the middle west, *Bull. Amer. Met. Soc.*, **35**, pp. 166—170.

MORGAN, G. W., 1956: On the wind-driven ocean circulation, *Tellus*, **8**, pp. 301—320.

MUNK, W. H., 1950: On the wind-driven ocean circulation, *J. Met.*, **7**, pp. 79—93.

NEUMANN, G., 1960: On the dynamical structure of the Gulf Stream, etc., *J. Geophys. Res.*, **65**, pp. 239—247.

NEWTON, C. W., 1959: Synoptic comparisons of jet stream and Gulf Stream systems, U. of Chicago, Dept. of Meteor. Tech. Rep. # 1 to ONR, Feb. 1959, 34 pages (Also in Rossby Memorial Volume, *The Atmosphere and the Sea in Motion*, Oxford Univ. Press, 1960.)

POLSON, R. K., 1958: Sunrise wind - How to find and use it, *Business and Commercial Aviation*, Feb. 1958, pp. 42—44.

STOMMEL, H., 1948: The westward intensification of wind-driven ocean currents, *Trans. Amer. Geophys. Union*, **29**, pp. 202—206.

— 1958: *The Gulf Stream*, U. of Calif. Press, Berkeley, Calif. 202 pages.

TEPPER, M., et al., 1954: *Pressure jump lines in the mid-western United States*, Res. Paper #37, U.S. Dept. of Commerce, Weather Bureau, 70 pages.

— 1955: On the generation of pressure-jump lines by the impulsive addition of momentum to simple current systems, *J. of Met.*, **12**, pp. 287—297.

— and NEWSTEIN, H., 1956: On the impulsive addition of momentum to simple current systems by asymmetrical velocity profiles, *J. of Met.*, **13**, pp. 314—315.

VERONIS, G., 1956: Partition of energy between geostrophic and non-geostrophic oceanic motions, *Deep-Sea Research*, **3**, pp. 157—177.

WAGNER, A., 1939: Über die Tageswinde in der freien Atmosphäre, *Gerl. Beitr. Geophys.*, **47**, pp. 172—202.

On the problem of truncation error

By M. G. WURTELE, University of California, Los Angeles

(Manuscript received March 28, 1961)

Abstract

It is argued that a more systematic study of truncation error is highly desirable. A number of integrations are commuted from the same partial differential equation using a variety of discretization schema. In particular, the difficulty of advecting sharp gradients of a conservative quantity is emphasized.

1. Introduction

This paper suggests the desirability of a systematic and, if possible, unified study of the finite-difference approximations to the partial differential equations commonly used in meteorological dynamics. Investigations of stability and accuracy of solutions of such difference equations are numerous in the literature, but most are studies of a particular approximation to a particular equation, with no attempt to contribute more general theorems, or even insights which might be applicable in other cases.

With a single exception, all examples computed in this paper are approximations to the one-dimensional linear conservative equation

$$\dot{\zeta} = \frac{\partial \zeta}{\partial t} + U\frac{\partial \zeta}{\partial x} = 0 \qquad U = \text{constant}, \qquad (1)$$

selected because of its very frequent occurrence in meteorological integrations. ζ may, of course, be any conservative quantity, but for convenience may be referred to as the vorticity. The solution of this equation is obviously

$$\zeta = f(x - Ut) \qquad (2)$$

with the initial condition

$$\zeta = f(x), \qquad t = 0$$

so that any initial ζ-pattern moves along with speed U without change in shape. It is the fact that the solution is exact for arbitrary initial conditions that is important here. The usual study of truncation error places emphasis upon the behavior of a single eigenfunction of a linear system, and although arbitrary initial states can be represented, the computations are seldom carried out. The techniques of the present paper are applied to the very important problem of forecasting the movement of sharp maxima and strong gradients. Such configurations are representable only by a large number of eigen-components, so that the conventional approach would require considerable computation. It will be noted that the accuracy of the results is well below the level one might anticipate from articles which emphasize only the stability of the computational schema and the preservation of integral invariants of the motion.

Boundary conditions are not considered, the solutions being valid on the infinite line. This seems justified, since the problems associated with boundaries have become less important with the larger grids used in most integrations.

It would be highly desirable to have com-

parable results in two space dimensions, but the techniques which would permit this are unknown to the author.

2. Differential-difference approximations

As a beginning, some insight is gained by the integration of differential-difference approximations, but owing to the sparcity of references to this technique in the meteorological literature, a few preliminary remarks are necessary. The use of differential-difference equations to approximate partial differential equations is in fact a classical procedure. A recent study by PINNEY (1958) contains a thirteen-page bibliography of references dating back to a paper by James Bernoulli in the early part of the 18th century. BATEMAN (1943) in a characteristically learned and elegant survey lists a variety of physical problems to which differential-difference equations have been applied. Among meteorologists, the first to bring the powerful classical results to bear on our current problems were apparently certain Soviet researchers, the most convenient reference in this connection being a paper by OBUKHOV (1957) which has been translated into English.

The simplest possible difference equation arises from the approximation of a derivative by a finite difference, say

$$\dot{u} \equiv \frac{du}{dt} = \frac{1}{2\tau}[u(t+\tau) - u(t-\tau)]. \qquad (3)$$

If this approximation is considered as an identity, what restrictions does this place on the function $u(t)$, or in other words, what class of functions is represented by all possible solutions of (3), treated as a differential-difference equation?

We note first that any second degree polynomial

$$u = a_0 + a_1 t + a_2 t^2$$

satisfies (1) identically. Further, an exponential

$$u = be^{st}$$

is also a solution provided that s is a root of the transcendental equation

$$\sinh s\tau = s\tau \qquad (4)$$

Thus the series

$$u(t) = a_0 + a_1 t + a_2 t^2 + \sum b_n e^{s_n t} \qquad (5)$$

is a general solution, where the s_n satisfy (4), and the b_n may be arbitrarily chosen. The class of functions which may be represented by such a series as (5) is an interesting and difficult mathematical problem. If the s_n are imaginary, this class corresponds to the almost-periodic functions; if they are real and negative, the series is a special case of a Dirichlet series. About these two classes of functions certain informative theorems exist, but in the present case, where the s_n are all complex, the problem is much more difficult and beyond the scope of this paper. Thus the simplest possible question one can ask concerning truncation of derivatives—what class of functions is represented by the differencing? — leads to profound and unsolved mathematical problems. For our purposes it is sufficient to note that if s_n is a solution of equation (4), so also is $-s_n$. Thus the series in (5) represents in general a function which is exponentially unbounded for $t \to \pm\infty$, although the original function u approximated by (3) may have been bounded over the entire t-axis. If half of the coefficients b_n are chosen zero, the solution u is then exponentially unbounded on one half-axis and exponentially vanishing on the other. If this type of result characterizes even the simplest finite difference approximation, it must be expected that in more complex finite-difference schemes, singularities and other undesirable properties of solutions may be introduced which are not present in the continuous representation of the infinitesimal calculus.

Another problem familiar to meteorologists is illustrated by a simple differential-difference equation: the dependence of the properties of the solution on the scheme of difference employed. If (3) is replaced by a forward difference,

$$\tau\dot{u}(t) = u(t+\tau) - u(t),$$

the equation corresponding to (4) is

$$1 + \tau s = e^{\tau s}$$

In this case there is only one root s_n with negative real part, all the rest having positive real parts. Thus a bounded function $u(t)$ can in general be represented only on the negative t-axis. The converse is true for solutions of the differential-difference equations arising from a

backward difference, for which the equation corresponding to (5) is

$$1 - \tau s = e^{-\tau s}$$

An infinity of roots s_n exist with negative real parts, a single root with positive real part. Thus we may generalize, admittedly rather vaguely, that as approximation to $\dot{u}$ the forward and backward differences are mirror images of each other; as one solution becomes unbounded, the other approaches zero, and vice versa. The centered difference is symmetric in $\pm t$, and therefore is the more suitable approximation for solutions which should oscillate or maintain constant amplitude. Stated otherwise, the centered difference is a second order difference and thus exhibits a symmetry not derivable with one-sides, first order differencing. Perhaps it is possible to establish precise general principles as to when each approximation is best, but this has not been done, even for such simple equations as the above.

The example studied by Obukhov is the one-dimensional conservation equation (1). It is noted that if a centered difference be taken for the x-derivative,

$$\frac{\partial \zeta_k}{\partial x} = \frac{1}{2h}(\zeta_{k+1} - \zeta_{k-1}),$$

where

$$\zeta_k = \zeta(kh, t), \qquad k = 0, \pm 1 \pm 2, \ldots$$

the conservation equation becomes

$$\frac{d}{d\tau}\zeta_k = \frac{1}{2}(\zeta_{k-1} - \zeta_{k+1}), \tag{6}$$

where $\tau = Ut/h$ is a non-dimensional time variable. The fundamental solution for this differential-difference equation is

$$\zeta_k(\tau) = J_k(\tau),$$

where J_k is a Bessel function of the first kind. Since this solution at the initial time $t = 0$ is zero for every value of k except $k = 0$, for which it is unity, the fundamental solution may be generalized by superposition:

$$\zeta_k(\tau) = \sum_{p=-\infty}^{\infty} a_p J_{k-p}(\tau). \tag{7}$$

This satisfies the arbitrary initial field

$$\zeta_k(0) = a_k.$$

The solution $J_{k-p}(\tau)$ thus acts as a Green's or influence function. In so far as the basic requirement of conservation of total vorticity is concerned, this solution is satisfactory. If we write it as

$$\zeta_k(\tau) = \sum_{q=-\infty}^{\infty} a_{k-q} J_q(\tau),$$

we have the total vorticity

$$\begin{aligned} Z(\tau) \equiv \sum_{k=-\infty}^{\infty} \zeta_k(\tau) &= \sum_{k=-\infty}^{\infty} \sum_{q=-\infty}^{\infty} a_{k-q} J_q(\tau) \\ &= \sum_{q=-\infty}^{\infty} J_q(\tau) \sum_{k=-\infty}^{\infty} a_{k-q} \\ &= Z_0 \sum_{q=-\infty}^{\infty} J_q(\tau), \end{aligned}$$

where Z_0 is the total initial vorticity. The well-known Laurent expansion for the Bessel coefficients is

$$e^{\frac{1}{2}\left(\alpha - \frac{1}{\alpha}\right)} = \sum_{q=-\infty}^{\infty} \alpha^q J_q(\tau),$$

and the result $Z \equiv Z_0$ follows upon setting $\alpha = 1$.

However, the vorticity *distribution* is not maintained as given initially, that is, the vorticity on each parcel is not advected with constant velocity, and it is this failure which was first pointed out by Obukhov. Obukhov computes the solution (7) for a given initial state represented by a unit step function at $t = 0$ and finds that a train of spurious waves follows the propagation of the disturbance. We may see this phenomenon, as well as another which suggests that Obukhov's conclusions are questionable, by taking the simpler initial field

$$\begin{aligned} \zeta(x, 0) &= 1, & |x| \leq 3/2 \\ &= 0 & \text{otherwise.} \end{aligned}$$

This corresponds to a single maximum of vorticity being advected with velocity U, the continuous solution being

$$\begin{aligned} \zeta(x, t) &= 1, & |x - Ut| \leq 3h/2 \\ &= 0, & \text{otherwise.} \end{aligned} \tag{7a}$$

The corresponding discrete initial field is

$$\begin{aligned} \zeta_k(0) &= a_k = 1, & k = -1, 0, 1 \\ &= 0 & \text{otherwise.} \end{aligned} \tag{7b}$$

It is evident that the discrete field (7 b) corresponds not only to the two step-functions (7 a)

but to any distribution which satisfies the discrete values. Thus (7 b) may be thought of as a general representation of a sharp maximum with steep gradients on either side. Since the continuous conservation equation (1) merely requires the initial pattern to move without change in shape, no difficulties are encountered in the true solution. For simplicity, the figures are all drawn for the particular continuous initial condition (7 a).

Thus the solution (7) takes the form

$$\zeta_k(\tau) = \sum_{p=-1}^{1} J_{k-p}(\tau). \tag{8}$$

It is to be emphasized that although the particular initial field chosen makes for clarity in the discussion of the solution, all remarks are quite general and apply equally to more complex initial conditons.

We first note that the speed of propagation of the disturbance is approximately correct, since $J_k(\tau)$ has a maximum for k approximately equal to τ. Next we evaluate the solution (8) far behind the disturbance after a long time, through use of the asymptotic expansion of the Bessel function for large argument:

$$J_k(\tau) \cong \left(\frac{2}{\pi\tau}\right)^{\frac{1}{2}} \cos\left[\tau - \left(k + \frac{1}{2}\right)\pi\right] \tag{9}$$

In the solution (8), substituting from (9), the first and third Bessel functions cancel, giving the simple asymptotic solution (9), $\zeta_k(\tau) \cong J_k(\tau)$. This is clearly oscillatory in both k and τ, an example of Obukhov's "parasitic waves". Note that the amplitude falls off like $\tau^{-1/3}$. The existence of the oscillations we may associate with use of the centered difference.

However, this asymptotic solution is valid only in the region $k \ll \tau$, well behind the disturbance. To study the propagation of the disturbance itself one must effect an expansion of $J_k(\tau)$ in the region $k \doteq \tau$. This result was derived by WATSON (1948, pp. 245—248):

$$J_k(\tau) \sim \frac{1}{3\pi} \sin\left(\frac{\pi}{3}\right)\left(-\frac{2}{3}\right)!\left(\frac{1}{6}\tau\right)^{-1/3}, \qquad \begin{matrix} |1 - k/\tau| \ll 1 \\ 6 \ll \tau \end{matrix}$$

For this value we obtain the result

$$\zeta_k(\tau) \cong 0.775\, \tau^{-1/3}$$

The region of maximum vorticity is advected with the proper speed and its amplitude, instead of remaining constant, falls off like $\tau^{-1/3}$.

Consider now the same equation with a backward difference for $\partial\zeta/\partial x$:

$$\frac{d\zeta_k}{d\tau} = \zeta_{k-1} - \zeta_k. \tag{10}$$

Obukhov shows that this is also an approximation for the advection equation with diffusion,

$$\frac{\partial\zeta}{\partial t} = -U\frac{\partial\zeta}{\partial x} + \alpha\frac{\partial^2\zeta}{\partial x^2}$$

with a centered difference in x and $\alpha = hU/2$. We may note that (10) is the differential equation for the Poisson process, a time-dependent stochastic process which assumes that a constant probability Pdt exists for the occurence of a single event in the time interval dt. In the equation (10), $\zeta_k(\tau)$ is interpreted as the probability that exactly k events occur in the time interval τ, with probability $P = 1$ that exactly one event occurs in the time interval dt. (See for example, FELLER, 1957, pp. 400—402.)

Stochastic processes and diffusion are, of course, intimately related, but it is a little difficult to see their connection with conservation equation. The fact is that although we are entirely familiar with the use of a single partial differential equation to represent a variety of physical phenomena, the introduction of finite differences immensely expands this variety. It presumably behooves meteorologists to become familiar also with these new representations if they are to employ finite differences with maximum effectiveness.

The solution of (10), namely the Poisson distribution, is well known:

$$\zeta_k(\tau) = \frac{e^{-\tau}\tau^k}{k!} \qquad k \geq 0$$
$$= 0 \qquad k < 0,$$

with the initial value

$$\zeta_k(0) = 1, \qquad k = 0$$
$$= 0, \qquad k \neq 0.$$

The corresponding influence function is

$$e^{-\tau}\tau^{k-p}/(k-p)!$$

so that the solution for the arbitrary initial field

$$\zeta_k(0) = a_k$$

is

$$\zeta_k(\tau) = \sum_{p=-\infty}^{k} a_p e^{-\tau} \frac{\tau^{k-p}}{(k-p)!}$$

$$= e^{-\tau} \sum_{q=0}^{\infty} a_{k-q} \frac{\tau^q}{q!}.$$

As before, it can be shown that the total vorticity is conserved, and that the speed of propagation is approximately correct, since $e^{-\tau}\tau^q/q!$ has a maximum for $\tau \doteq q$.

For the single initial maximum of vorticity (7 b) this becomes

$$\zeta_k(\tau) = e^{-\tau}\left[\frac{\tau^{k-1}}{(k-1)!} + \frac{\tau^k}{k!} + \frac{\tau^{k+1}}{(k+1)!}\right]$$

$$= \frac{e^{-\tau}\tau^{k+1}}{(k+1)!}\left[1 + \frac{k+1}{\tau} + \frac{k(k+1)}{\tau^2}\right]. \qquad (11)$$

As might have been expected with the use of a backward difference, the solution is monotonic in x, thus eliminating the "parasitic waves". On this basis, Obukhov finds it superior to the solution (8). However, it is equally important to note that for large τ the asymptotic value is $\frac{3e^{-\tau}\tau^{k+1}}{(k+1)!}$. This is $< \tau^{-1/2}$ where $k = \tau$, in contrast to $\tau^{-1/3}$ of the centered-difference solution. As a comparison we may compute the *maximum* value of the two solutions. For a wind speed $U = 20$ m/sec and a grid distance $h = 300$ km, the value $\tau = 6$ corresponds to 9.0×10^4 seconds, or a little more than one day. At the end of this time we have (from Poisson distribution tables)

$$[\zeta_k(6)]_{\max} \sim 0.48 \qquad \text{(backward difference)}$$

and (from tables of Bessel functions)

$$[\zeta_k(6)]_{\max} \sim 0.86 \qquad \text{(centered difference)}$$

Instead of conserving and advecting the vorticity maximum strictly with the velocity U, the centered difference approximation has dispersed a small portion of it into the spurious waves, but the backward-difference approximation has diffused away 52 per cent, both up- and down-stream. The exact solutions for

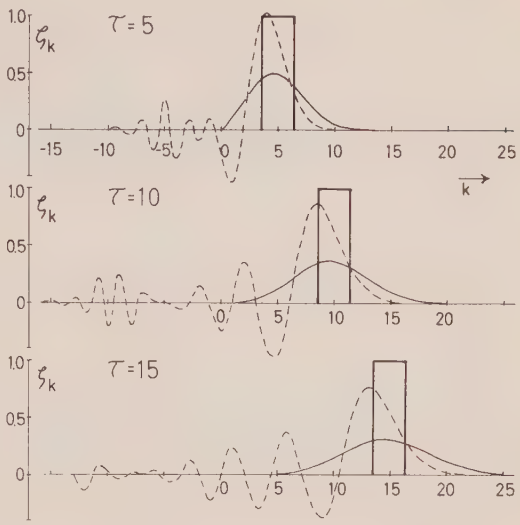


Fig. 1. Three solutions of the advection equation for (non-dimensional) times $\tau = 5$, 10, 15.
——— (exact) solution of continuous equation (4)
— — — — solution (8) of centered differential-difference equation
——— solution (11) of backward differential-difference equation
For typical meteorological values, ten units of non-dimensional time correspond to about 42 hours.

three different times, $\tau = 5$ (about 21 hours), $\tau = 10$, and $\tau = 15$ are represented in Figure 1. Curves have been sketched through the discrete points as an aid to perspicuity.

A glance suffices to show that neither finite-difference solution is a satisfactory approximation to the continuous solution. In Obukhov's example, the effect of the diffusion on reducing the maximum vorticity was not apparent, so that the choice between the differencing schemes seemed easier. Perhaps a more extended comparison is indicated, using more realistic wave-patterns.

The most frequently integrated equation of meteorology, the barotropic vorticity equation, expresses a conservation law, the final output being not the vorticity but the height of the pressure surface. The height is a twice-integrated function of the vorticity field, and as such may be expected to be a much smoother field.

This smoothing is particularly effective in eliminating the spurious oscillations contained in the solution (8). To see this it is convenient to take an initial vorticity pattern representing

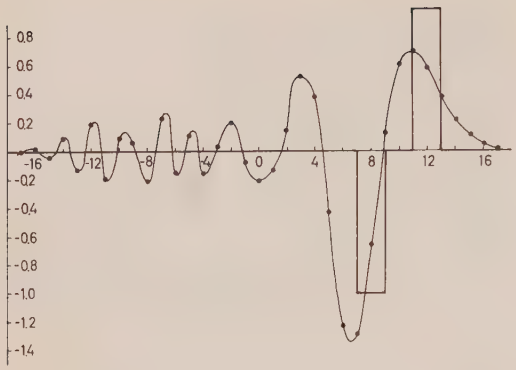


Fig. 2 a. Solution of the centered differential-difference equation contrasted with exact solution. The "vorticity" pattern.

exact solution

differential-difference solution

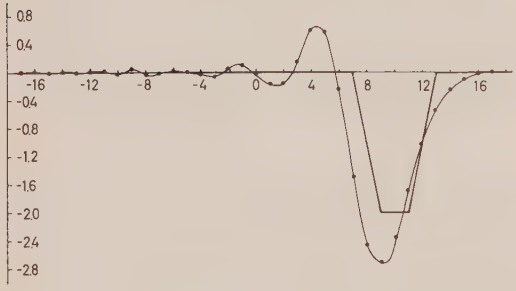


Fig. 2 b. The integrals of the solutions of Figure 2 a. The "velocity" pattern.

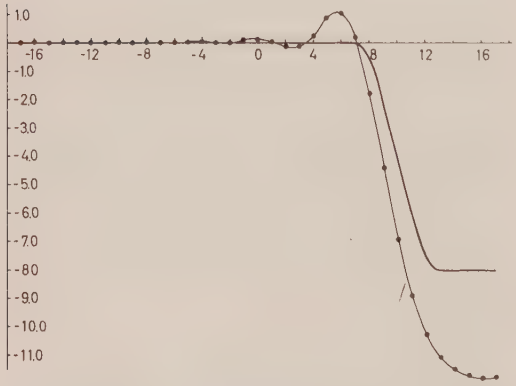


Fig. 2 c. The integrals of the function of Figure 2 b. The "height" pattern.

a sharp minimum. The solution was then integrated according to the simple formula

$$V_k = \sum_{p=\infty}^{k} \zeta_p,$$

and this in turn was integrated,

$$\psi_k = \sum_{p=-\infty}^{k} V_p.$$

The ζ, V, and ψ fields are plotted in Figures 2 a, b, c. As expected, the accuracy of the approximation increases markedly. If the finite difference approximation produces oscillatory errors, it is fortunate to be primarily interested in quantities which are space integrals of the quantities governed by the dynamic equations.

3. An application to the wave equation

To emphasize the importance of choosing a finite-difference scheme which conserves conservative quantities, we may examine an approximation to the wave equation,

$$\frac{\partial^2 \xi}{\partial t^2} = c^2 \frac{\partial^2 \xi}{\partial x^2}. \tag{12}$$

Since the wave equation is derivable from two equations similar to the advection equation,

$$\frac{\partial \xi}{\partial t} = c \frac{\partial \eta}{\partial x} \qquad \frac{\partial \eta}{\partial t} = c \frac{\partial \xi}{\partial x}$$

by elimination of either dependent variable, it is natural to expect solutions similar to the ones with which we have been dealing. This is the case. A centred difference for x yields

$$\frac{d^3 \xi_k}{d\tau^2} = \xi_{k+1} + \xi_{k-1} - 2\xi_k, \tag{13}$$

where $\tau = ct/h$, h being the x-difference as before. The fundamental solution of (13) is

$$\zeta_k(\tau) = J_{2k}(2t),$$

but the general solution is complicated by the fact that if the medium is unbounded in the x-direction, two initial fields are required to specify a unique solution, both $\dot{\xi}$ and its first time derivative. The corresponding solution of the wave equation (that of D'Alembert) is

$$\xi(x,t) = \frac{1}{2}f(x-ct) + \frac{1}{2}f(x+ct) + \frac{1}{2}\int_{x-ct}^{x+ct} g(s)\,ds \qquad (14)$$

where $\xi(x,0) = f(x)$ may be interpreted as the initial displacement field and $\dot{\xi}(x,0) = g(x)$ as the initial velocity field. If we assign the arbitrary conditions

$$\xi_k(0) = a_k$$
$$\dot{\xi}_k(0) = b_k,$$

the required solution of the differential-difference equation (13) is obtained as before by use of an influence function:

$$\xi_k(\tau) = \sum_{p=-\infty}^{\infty} a_p\, J_{2(k-p)}(2\tau) + \sum_{p=-\infty}^{\infty} b_p \int_0^{\tau} J_{2(k-p)}(2s)\,ds. \qquad (15)$$

This solution has apparently been known for half a century. It is attributed by BATEMAN to KOPPE (1901) in an obscure reference not easily available. The first term on the right hand side behaves like the advective solution (7), except as is easily seen, since $J_{2n}(\tau) = J_{-2n}(\tau)$ if a_k and b_k are even functions of k, then ξ_k is also an even function of k, This represents the propagation of the wave in both directions, corresponding to the continuous solution (14). It is however, the second term that concerns us now. Koppe noted that for large values of the time τ[1],

$$\int_0^{\tau} J_{2q}(2s)\,ds \cong 1,$$

which means that whereas the initial displacements a_k die out like $\tau^{-1/3}$, the initial velocities b_k are conserved in magnitude. From the point of view of the meteorologist this is a little frightening. A quantity such as the height of a constant pressure surface is much more easily and accurately measured than the vertical velocity of the parcels on the pressure surface. Thus it would seem possible for the solution after a long time to contain little other than error, the accurate displacement observations having been dispersed into spurious waves and the erroneous initial velocities carefully preserved.

[1] See Appendix for a proof of this expansion.

It is difficult to see how this source of truncation error could be eliminated. That it is a property of the finite differencing, not of the continuous solution, can be shown. Let us consider an example. The initial displacement field we maintain as before:

$$f(x) = 1 \qquad |x| \leqq 3/2$$
$$= 0 \qquad \text{otherwise,}$$

with the corresponding discrete field

$$a_k = 1 \qquad k = -1, 0, 1$$
$$= 0 \qquad \text{otherwise.}$$

Now let the initial velocity field be identically zero, except in a small region of width 2ε at the origin:

$$g(x) = E \qquad |x| \leqq \varepsilon$$
$$= 0 \qquad \text{otherwise.}$$

This may be interpreted as an instrumental error of magnitude E in the instrument or instruments measuring velocity at the observing station at the origin (assuming these stations to be evenly distributed at $k = 0, \pm 1, \pm 1, \pm 2, \ldots$):

$$b_k = E \qquad k = 0$$
$$= 0 \qquad \text{otherwise.}$$

D'Alembert's solution (14) gives for the displacement term (when $ct > 2$)

$$f(x-ct) + f(x+ct) = 1 \qquad |x \pm ct| \leqq 3/2$$
$$= 0 \qquad \text{otherwise;} \qquad (16a)$$

and for the velocity term

$$\frac{1}{2}\int_{x+ct}^{x+ct} g(s)\,ds = \varepsilon E \qquad \varepsilon - ct < x < ct - \varepsilon$$
$$= 0 \qquad \text{otherwise,} \qquad (16b)$$

The continuous solution thus propagates an error proportional to ε; so that isolated errors do not contribute appreciably in the integration. But the discrete system is at the mercy of the errors of the isolated observing stations, and here the least accurate initial conditions are asymptotically dominating.

The solution (15) is easily computed for the above initial conditions. Since k takes only

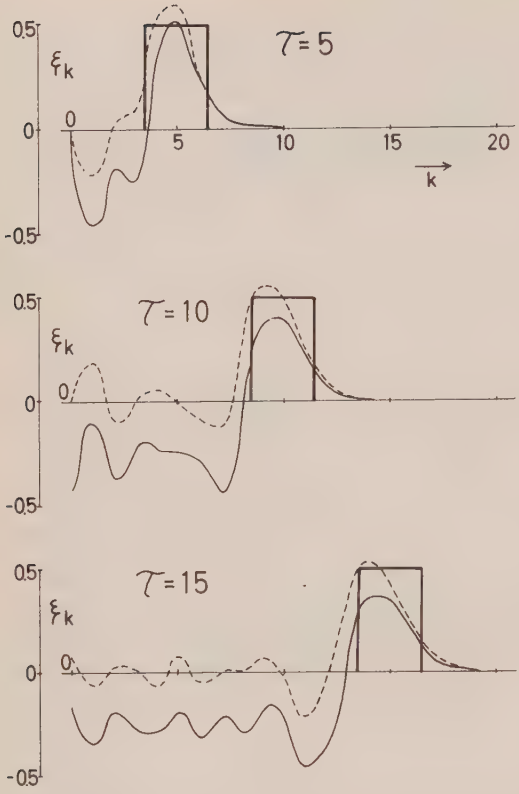


Fig. 3. Solutions of the wave equation for (non-dimensional times $\tau = 5, 10, 15$.
——— (exact) solution (16) of continuous equation
— — — — solution (17)of differential-difference equation without initial error
——— solution (17) of differential-difference equation with initial error at single isolated point $k = 0$.

integral values we have (WATSON, 1948, p. 545)

$$\int_0^{\tau} J_{2(k-p)}(2s)\,ds = \Big[\sum_{q=0}^{\infty} J_{2(k-p)+2q+1}(s)\Big]_0^{2\tau}$$

$$= \sum_{q=0}^{\infty} J_{2(k-p+q)+1}(2\tau).$$

The final solution is then

$$\xi_\tau(\tau) = \sum_{q=k-1}^{k+1} J_{2q}(2\tau) + E\sum_{q=0}^{\infty} J_{2(k+q)+1}(2\tau). \tag{17}$$

The continuous solution (16 a, b) and the solution (17) for $E=0$ and for $E={}^1/_2$ are plotted in Figure 3. The increasing relative contribution of the error E is apparent. In systems of equations such as those of atmospheric motion, which contain both conservation of some quantity and wave-propagation of another, this effect could be disastrous. It is a little difficult to say precisely what time scale is involved here. For the long external gravity waves with phase speeds of hundreds of meters per second, the time scale of, say, 10 non-dimensional units is only a few hours. If internal gravity waves are considered, with phase speeds of the order of the wind speed, the time scale is comparable to that of Figure 1.

4. Partial difference formulations

The more conventional approach to partial differential equations in meteorological computations is to approximate all derivatives by finite differences, a method yielding apparently superior results to that of the previous sections. It is well known that if the conservation equation (1) is approximated by a forward difference in time and a centered difference in x, the solution is unbounded as time increases. However, an interesting result is obtained by taking one-sided differences for each derivative forward in time and backward in space. If $x = k\,\Delta x$, $t = n\,\Delta t$, we have

$$\frac{\zeta_{k,n+1} - \zeta_{k,n}}{\Delta t} = -U\frac{\zeta_{k,n} - \zeta_{k-1,n}}{\Delta x}$$

or

$$\zeta_{k,n+1} = (1-\alpha)\zeta_{k,n} + \alpha\zeta_{k-1,n} \tag{18}$$

where $\alpha = U\,\Delta t/\Delta x$. As in example (10) of Section 2, in which a onesided derivative was used, the equation (18) may be given a probabilistic interpretation. If α is the probability of winning one unit at each trial, and $(1-\alpha)$ is the probability of winning no units, then $\zeta_{k,n}$ is the probability of possessing k units after n independent trials. It is to be expected, then, that the solutions of (18), viewed as an approximation to (1), will be bounded and stable for $0 < \alpha < 1$. This is easily shown to be true by conventional stability analysis. Further, since a probability cannot be negative, no unwanted oscillations will be present.

The classical solution to this Bernoulli problem is (e.g. FELLER, 1957, pp. 136—137)

$$\zeta_{k,n} = \binom{n}{k}(1-\alpha)^{n-k}\alpha^k,$$

which may be generalized for arbitrary initial conditions (that is, say, for various initial capitals of the gambler). If

$$\zeta_{k,0} = a_k$$

we have

$$\zeta_{k,n} = \sum_{j=-p}^{p} a_{-j}\binom{n}{k+j}(1-\alpha)^{n-k-j}\alpha^{1},$$

where p may be made as large as desired. For the particular initial condition (7b) the solution is

$$\zeta_{k,n} = \sum_{j=-1}^{1}\binom{n}{k+j}(1-\alpha)^{n-k-j}\alpha^{j}. \qquad (19)$$

In a study of a similar solution in BLAIR et al. (1957), posthumously published (Appendix 3, by von Neumann, pp. 77—81), it is pointed out that for large n the DeMoivre-Laplace limit theorem is applicable here. The $\zeta_{k,n}$ of equation (19) may be interpreted as the probability that the number of wins lies between $k-1$ and $k+1$, and the solution is approximated by the normal distribution function (FELLER, 1957, p. 172)

$$\zeta_{k,n} \cong \Phi\left\{\frac{k+\frac{3}{2}-\alpha n}{[2n\alpha(1-\alpha)]^{1/2}}\right\} - \Phi\left\{\frac{k-\frac{3}{2}-\alpha n}{[2n\alpha(1-\alpha)]^{1/2}}\right\}, \qquad (20)$$

where

$$\Phi(x) = (2\pi)^{-1/2}\int_{-\infty}^{x} e^{-y^2/2}dy.$$

It is evident that this solution $\zeta_{k,n}$ has a maximum in the neighborhood of $k = n\alpha$, that is at $k\,\Delta x = U\,\Delta t$. The speed of propagation of the disturbance is thus correct. However, as von Neumann pointed out, the factor $[2n\alpha(1-\alpha)]^{1/2}$ is a measure of the dispersion of a sharp maximum after n time intervals, or a real time interval proportional to $n\alpha$. Therefore, the spurious diffusion is greater for α small.

These results seem consistent with those of Section 2 in which a one-sided differential-difference equation, amenable to interpretation as representing a stochastic process, diffused away some of the sharp maximum of the initial state. The analogy seems even closer if we regard equation (18) as representing a sort of random walk, with α the probability of a unit-step forward and $(1-\alpha)$ the probability of remaining stationary. This is evident if (18) is compared with the conventional random walk equation (e.g. FELLER, 1957, p. 318). Whereas the conventional random walk is, in the limit, a classical diffusion process, the solution of equation (18) is a forward translation and a backward diffusion.

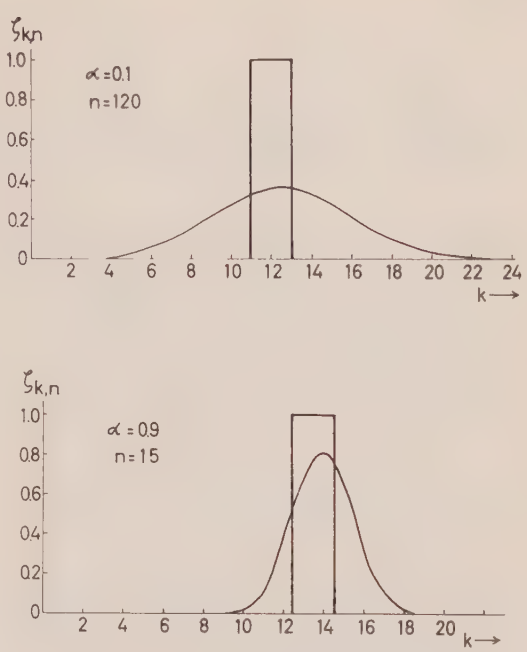


Fig. 4. Approximations according to (14) of solutions to the conservation equation with a forward time difference and backward space difference (18). (a) $\alpha = 0.1$, $n = 120$; (b) $\alpha = 0.9$, $n = 12$.

Two examples have been computed from (20) and are pictured in Figure 4. If U is taken as 20 m/sec and Δx as 300 km, then time intervals Δt of twenty-five minutes and 225 minutes, respectively correspond to $\alpha = 0.1$ and 0.9. In order to make the elapsed time comparable with the cases $\tau = 10$ in the previous figures (about 42 hours), the respective values $n = 120$ and $n = 12$ were selected. The marked difference in the extent of the diffusion is apparent. The parameter α must be less than but close to unity if the solution of the difference equation is to be a satisfactory approximation to the solution of the conversation equation.

We may now consider the more usual approximation to equation (1), which uses

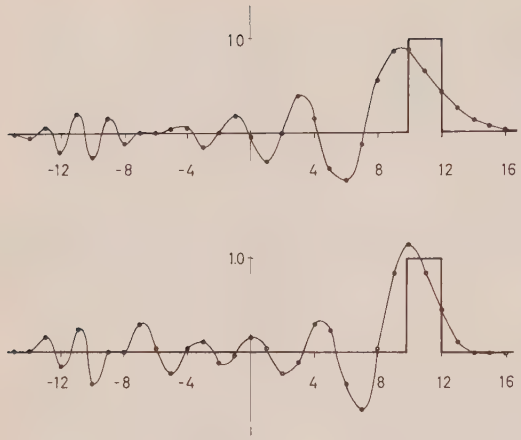


Fig. 5. Computed solutions of (21) with a forward difference for the first step. (a) $\alpha = 0.5$, $n = 23$; (b) $\alpha = 0.9$, $n = 13$.

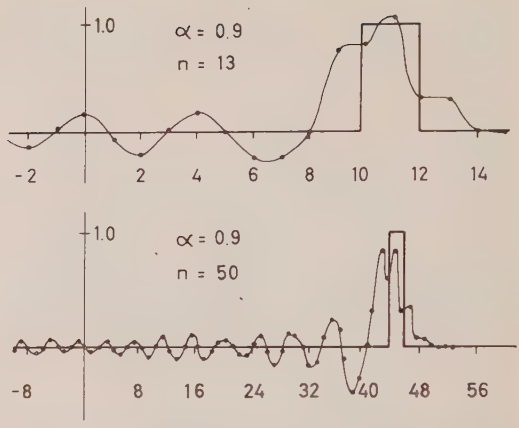


Fig. 7. Computed solution of (21), assigning true values at the first time step as a second initial condition. (a) $\alpha = 0.9$, $n = 13$. Cf. Figure 5 b. (b) $\alpha = 0.9$, $n = 50$. Cf. Figure 6 b.

centered differences for both time and space. The second-order partial difference equation is of the familiar form

$$\zeta_{k,n+1} = \zeta_{k,n-1} - \alpha(\zeta_{k+1,n} - \zeta_{k-1,n}), \quad (21)$$

where α has the same definition as before. As is well known, the solution is stable for $\alpha < 1$. A number of problems in connection with (21) have been discussed by GATES (1959, 1960), to which detailed reference will be made.

As with the centered differential-difference formulation, the solution admits both positive and negative values, and thus cannot represent the probability of a stochastic event. In fact, (21) is formally the meaningless special case of the equation governing the stochastic birth-and-death process (FELLER, 1957, p. 407) if the probability of death is set equal to —1.

The fact that (21) is of an order higher than (18) in the time derivative requires an extra initial condition. The procedure used in numerical forecasting, dictated by the realities of synoptic observtions, has been to apply a forward time difference for the first time step, and to use equation (21) from the second step onward. Another possibility available to us here is to assign true values for the first time step as a second initial condition.

Unfortunately, an analytical expression for the solution of (21) is not possible without more computation than is required for a direct stepwise integration, and therefore the latter procedure has been followed. In Figure 5 are pictured the solutions corresponding to a total elapsed time comparable to previous examples, for the cases $\alpha = 0.5$ and $\alpha = 0.9$. In Figure 6 are the same solutions after almost four times as many time steps. Both of these computations used the conventional forward difference for the first time step. As an example of the result of using assigned true values for a second initial condition at the first step, Figure 7 pictures the solution corresponding to $\alpha = 0.9$, $n = 13$.

The spurious oscillations characterizing all solutions have been analyzed in detail by GATES (1959, pp. 12—15). The sharp maximum

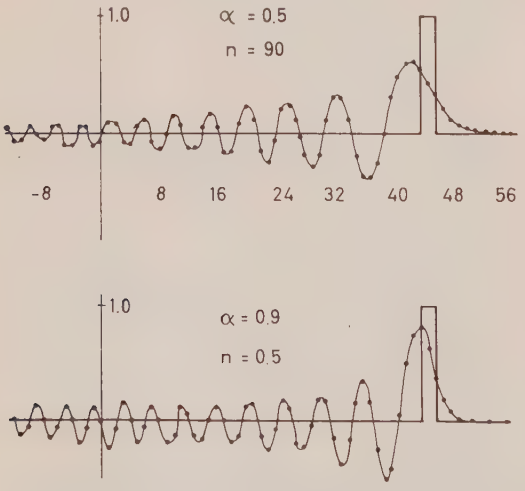


Fig. 6. The same as Figure 5 for (a) $\alpha = 0.5$, $n = 90$; $\alpha = 0.9$, $n = 50$.

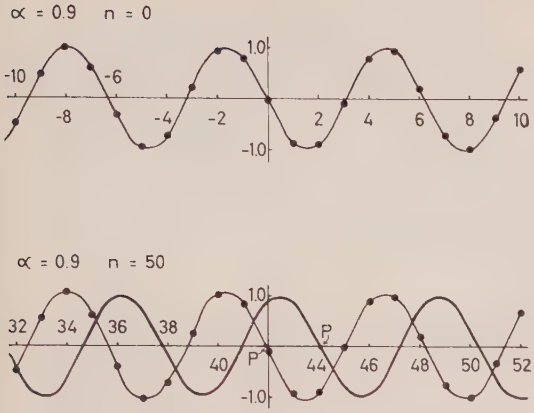


Fig. 8. The propagation according to (21) of an initially sinusoidally-distributed field. (a) $n = 0$. (b) $n = 50$; true solution, heavy line; computed solution, light line. The zeros labeled P and P' coincided initially.

of the present examples-although by no means unrealistic meteorologically, given conventional values of the grid distance—is composed of a large number of eigen-solutions of appreciable amplitude, and the spurious oscillations distort the solution much more seriously than would be the case with a single sine wave of wave length comparable in scale to that of the sharp maximum. It should be noted that this distortion is not appreciably reduced either by varying α nor by taking a second initial condition in place of the forward difference.

The satsifactory character of computations according to (21) when applied to a single eigen-function may be illustrated by a field the initial distribution of which is exactly sinusoidal. The boundary conditions become immediately operative in this case; but the results up to 50 time steps can be consulted at the center of the pattern, as yet uncontaminated by boundary effects. Figure 8 shows such a section. At five time steps the distortion in amplitude and phase speed would be still negligible on the scale of the diagram, and after 50 time steps the computation remains quite satisfactory, although the computed wave has fallen a little more than one-quarter wave length behind the true one.

The result is consistent with the analytic formulas derived by GATES (1959, pp. 12—13, 19—21) for the wave selected.

We may gain further insight by inspection of the actual Fourier series for a periodic

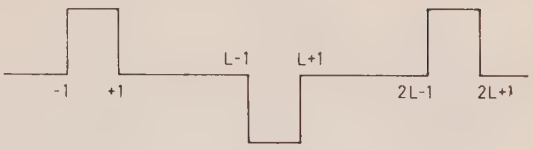


Fig. 9.

pattern of sharp maxima and minima, arranged as in Figure 9, in which the extremal regions are two units in width and are separated by L units. The series for this distribution is

$$\zeta(x) = \frac{4}{\pi} \sum_{m=0}^{\infty} \frac{1}{2m+1} \sin\left[\frac{(2m+1)\pi}{L}\right] \cos\left[\frac{(2m+1)\pi}{L} x\right].$$

Now consider the case in which the separation L becomes large compared to the width of the maxima. For, say, $L = 20$, certainly not an unusual separation of sharp gradients on a cyclonic scale, the amplitude of the wave of length $L/5$ is about two-thirds that of the wave of length L, and the amplitude of the wave of length $L/10$ is one-third of the latter value. These short waves are approaching grid-mesh dimensions, and accuracy in their numerical treatment is not to be expected.

Mr. G. E. Birchfield of the University of Chicago has kindly allowed me to exhibit a computation (unpublished) of his which contains a spurious oscillation, apparently due to this effect. The computation was a two-dimensional conservation of vorticity, using centered time and space differences, from a hurricane forecast. Figure 10 represents the distribution

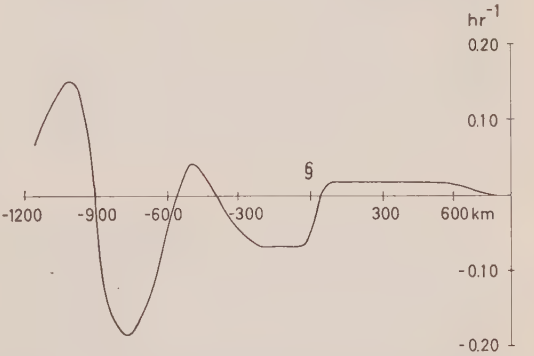


Fig. 10. Forecast vorticity distribution along direction of motion of a hurricane (after G. E. Birchfield) exhibiting spurious oscillations.

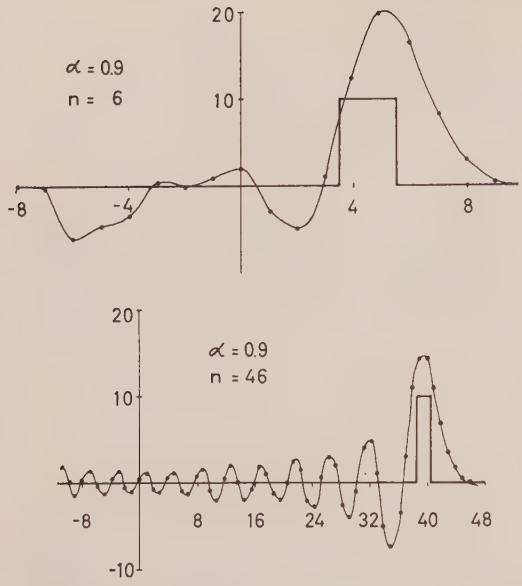


Fig. 11. Computed solutions of (21) using special starting procedure of Gates. (a) $n = 6$; (b) $n = 46$.

of vorticity along a line which is roughly the direction of movement of the system.

Gates (1960, pp. 4—12) has suggested a technique for reduction of this contaminating effect which calls for a special starting procedure. The first forward time difference is taken over a smaller interval $\Delta t/4$ the second step (a centered difference) over interval $\Delta t/2$, the third over Δt, and the fourth and all subsequent steps over $2\Delta t$. Applied to the example of the present paper, the results are pictured in Figure 11. The accuracy is not appreciably greater than that achieved by the previous scheme. Gates' conclusions are, of course, unexceptionable, but we see again the effect of a pattern which incorporates a large number of Fourier components of appreciable amplitude.

Acknowledgements

My thanks are due Mr. G. E. Birchfield for permission to draft and publish Figure 8 from his output sheets.

In the computations and preparations of figures, I was assisted by Messrs. R. P. Emig, II, Brian Finke, and Hugh Stone. Programming was the work of Mr. Finke.

The computations were performed on the facilities of the Western Data Processing Center, Graduate School of Business Administration, University of California, Los Angeles, the cooperation of which is gratefully acknowledged.

This research was performed with the support of contracts Nonr-233-21 and DA-04-495-ORD-1845.

Appendix

Let

$$I \equiv \int_0^\tau J_{2k}(2s)\,ds \qquad \text{(i)}$$

To obtain the asymptotic expansion for large τ, we express the Bessel function as an integral,

$$J_{2k}(2s) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \exp.\, 2i(s\sin\varphi - k\varphi)\,d\varphi$$

substitute into (i), setting $\sigma = s/\tau$, and obtain

$$I = \frac{\tau}{2\pi}\int_0^1\int_{-\pi}^{\pi} \exp.\, 2if(\sigma,\varphi)\,d\varphi d\sigma$$

where

$$f(\sigma,\varphi) = \tau\sigma\sin\varphi - k\varphi$$

We now apply the method of steepest descent:

$$\frac{\partial f}{\partial\sigma} = \tau\sin\varphi = 0 \qquad \text{for } \varphi = \varphi_0 = 0$$

$$\frac{\partial f}{\partial\varphi} = \tau\sigma\cos\varphi - k = 0 \text{ for } \sigma = \sigma_0 = k/\tau\cos\varphi_0 = k/\tau$$

Thus for $0 < k < \tau$ the saddle-point is on the path of integration, and for $\tau^{1/2} \gg 1$,

$$I \sim \frac{\tau}{2\pi}\int_0^1\int_{-\pi}^{\pi} \exp\, 2i\tau(\sigma - \sigma_0)\varphi d\varphi d\sigma$$

$$\sim \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp.\, 2i\sigma'\varphi' d\varphi' d\sigma'$$

$$\sim \frac{1}{\pi}\left(\int_{-\infty}^{\infty} \exp.\, iu^2 du\right)^2$$

$$\sim \frac{1}{\pi}\left(\sqrt{\pi}\right)^2$$

$$\sim 1$$

For $-\tau < k < 0$, the saddle-point $\varphi_0 = \pi$ may be used to obtain the same result. As $|k| \to \tau$ the value of the integral begins to fall off rapidly, and for $|k| > \tau$, the integral receives no contribution from the neighborhood of the saddle-point, so that its order of magnitude is $k^{-1/2}$.

REFERENCES

Bateman, H., 1943: "Some simple differential difference equations and the related functions," *Bull. Am. Math. Soc.*, **49**, 494—512.

Blair, N., Metropolis, N., Neumann, J. von, Taub, A. Tsingon, M., 1957: *A Study of numerical solution to a two-dimensional hydrodynamics problem*, LA-2165, Los Alamos Scientific Laboratory.

Feller, W., 1957: *An introduction to probability theroy and its applications*, volume one, 2nd edn., John Wiley and Sons, 461 pp.

Gates, W. L., 1959: *On the truncation error, stability and convergence of difference solutions of the barotropic vorticity equation*, Scientific Report No. 1, Contract AF 19(604)—3886, Department of Meteorology, University of California, Los Angeles.

— 1960: *Analytical studies of numerical error in dynamical weather prediction*, Final Report, Contract AF 19(604)—4965, Department of Meteorology, University of California, Los Angeles.

Koppe, M., 1901: *Die Ausbreitung einer Erschütterung an der Wellenmaschine durch einen neuen Grenzfall der Besselschen Funktionen*, Program Andreas-Real Gymnasium, Berlin, No. 96, 28 pp.

Obukhov, A. M., 1957: "On the accuracy of the prediction of advective changes in numerical forecasting," *Jour. Acad. Sci. USSR*, Geophysics Series, no. **9**, 1133—1141. Translated in *Meteorological Translations*, no. **1**, January 1959, Meteorological Branch, Department of Transport, Canada.

Pinney, E., 1958: *Ordinary difference-differential equations*, University of California Press, 262 pp.

Watson, G. N., 1948: *A treatise on the theory of Bessel functions*, 2nd edn., Cambridge University Press, 804 pp.

The Effect of Meanders on the Kinetic Energy Balance of the Gulf Stream

By FERRIS WEBSTER, Woods Hole Oceanographic Institution[1], Woods Hole, Massachusetts

(Manuscript received March 1, revised version June 28, 1961)

Abstract

Calculations have been made, using surface velocity observations, of the horizontal transfer of kinetic energy from meanders to mean flow at two separate localities in the Gulf Stream System. In both cases, it was found that there was a surface flow of kinetic energy from the meanders to the mean flow, exactly opposite to what would be expected if the meanders were frictionally driven. The observations suggest that the meanders derive energy from sources other than the kinetic energy of the mean flow. The role of the meanders in relation to the mean Gulf Stream appears analogous to that of large-scale eddies in relation to the general circulation of the atmosphere.

1. Introduction

Calculations have been made, for two separate regions, of the horizontal surface transfer of momentum by meanders in the Florida Current section of the Gulf Stream System. In both regions, it was found that the meanders transferred momentum against the velocity gradient, exactly opposite to what would be expected if they were frictionally driven. Or, in other words, there was at the surface, a net transfer of kinetic energy from the meanders to the mean flow.

Since no other such observations have been made of the transfer of kinetic energy in ocean currents, it has not been possible to determine the source of kinetic energy of meanders. It has been supposed by some (e.g.: ROSSBY, 1936, p. 6; STOMMEL, 1958, p. 107; VON ARX, 1954) that the meanders draw their energy from the kinetic energy of the mean flow and represent a mechanism for frictional dissipation of the mean flow. The observations here suggest that the opposite is true: that the mean flow is enhanced by the kinetic energy of meanders, and that the meanders should, therefore, derive their kinetic energy from sources other than the kinetic energy of the mean flow.

2. Observations

The data used to calculate eddy momentum fluxes were obtained from two separate sections across the axis of the Gulf Stream. One section is located off Onslow Bay, North Carolina, near Cape Hatteras, and the other is located in the Straits of Florida, running across the channel between Miami and Gun Cay, Bahamas. Figure 1 shows the location of these sections and their relation to the mean surface

[1] Contribution No. 1174 from the Woods Hole Oceanographic Institution.

Fig. 1. Florida Current region.

axis of the Gulf Stream, as estimated by the United States Coast and Geodetic Survey.

GEK data. The necessary velocity measurements were made with the geomagnetic electrokinetograph (GEK), first described by VON ARX (1950). These measurements are usually less than the true surface velocity, depending mainly on the depth of moving water in relation to the total depth of water (LONGUET-HIGGINS, STERN, and STOMMEL, 1954). It is general practice to determine the reduction of GEK velocities for any region by calculating the average ratio between surface velocities measured by dead-reckoning methods and surface velocities measured with the GEK. This ratio (called "k") has been determined for the Onslow Bay area from several hundred measurements by VON ARX, BUMPUS, and RICHARDSON (1955), who found for k, 1.46 ± 0.09. In the Straits of Florida it has been calculated by HELA and WAGNER (1954) to be 1.68 ±0.30. That the Straits of Florida have a somewhat higher k-value than does Onslow Bay is a result of the shallower depth of water in the Straits. All GEK observations presented here have been corrected by the appropriate value of k.

Onslow Bay. The data off Onslow Bay were collected by W. S. von ARX, D. F. BUMPUS, and C. G. DAY on Cruise 18 of the research vessel CRAWFORD during May and June, 1958. The structure of the meanders observed during this period has been described by WEBSTER (1961). The CRAWFORD sailed back and forth along a single line at right angles to the mean axis of the Stream. An observation of the surface water velocity was made each hour by means of the GEK. During the twenty-eight days of the cruise, 620 separate measurements, or fixes, of the surface water velocity were made. Between fixes, the component of surface velocity at right angles to the course of the ship was recorded continuously. Since the ship's track was across the mean axis, this component was equivalent to the downstream surface velocity. For this portion of the Stream, "downstream" means in the direction 040°T.

The time-average downstream surface velocity ($\bar{v}$) profile determined from CRAWFORD Cruise 18 data has a cyclonic shear of approximately 4.4×10^{-5} sec^{-1}, and an anticyclonic shear of approximately 3.0×10^{-5} sec^{-1}. The Coriolis parameter at this latitude has a value of 8.0×10^{-5} sec^{-1}. Because the averages of the cross-stream surface velocity ($\bar{u}$) are generally less than their standard deviations, the averages are not significantly non-zero. If there are mean cross-stream motions, they are less than 4 cm/sec.

Florida Straits. Between 1952 and 1958, the Marine Laboratory of the University of Miami was engaged in part of a long-term program to determine the characteristics of the Florida Current as it flows through the channel between Fowey Rocks, Florida (near Miami), and Gun Cay, Bahamas. The channel at this point is about 43 nautical miles (80 kilometers) wide. A total of 632 GEK observations of velocity in the surface layer from 42 of the cruises made by the Marine Laboratory were available for analysis (University of Miami, 1954, 1955, 1957, 1958).

The region surveyed by the University of Miami was divided into longitudinal zones. Only the observations made between latitudes 25°30′N and 25°59′N were used in calculating the averages for each zone. The zones and the region are shown in Figure 2. The limits of the zones are as follows:

Zone	From	To	Width
1	80°05′.5 W	80°01′.5 W	4′.0
2	01′.4	79°58′.5	3′.0
3	79°58′.4	55′.5	3′.0
4	55′.4	52′.5	3′.0
5	52′.4		3′.0
		49′.5	
6	49′.4	44′.5	5′.0
7	44′.4	39′.5	5′.0
8	39′.4	34′.5	5′.0
9	34′.4	29′.5	5′.0
10	29′.4	24′,5	5′.0
11	24′.4	19′.5	5′.0

The western boundary of the channel, Fowey Rocks, is at 80°06′.5 W, and the eastern boundary, Gun Cay, is at 79°18′.5 W, so that the zones chosen cover nearly the entire channel, and include all the observations available. In the Straits of Florida region, one minute of longitude is equal to 0.911 nautical miles (1.69 km). Therefore the width of the widest zones is 4.55 nautical miles (8.45 km).

The averages of the northward component of surface velocity (v) and the eastward component (u) were calculated for each zone

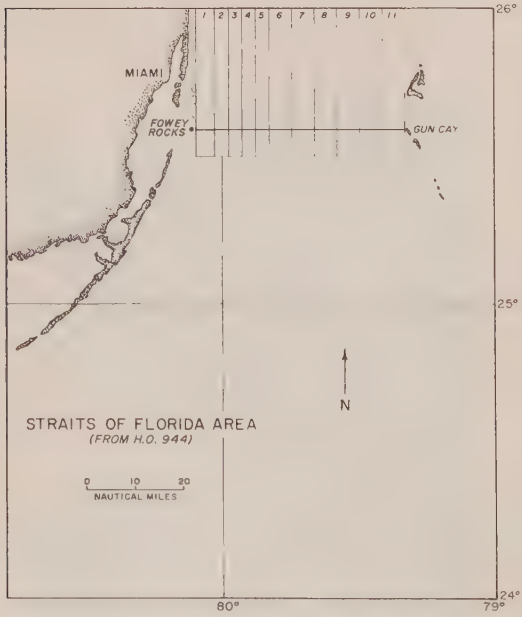


Fig. 2. Straits of Florida area, showing zones used to separate observations.

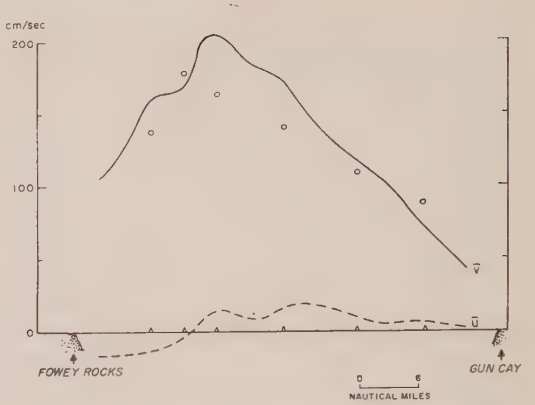


Fig. 3. Average velocity profile, Straits of Florida. Solid line is average downstream (north) velocity, dotted line is average cross-stream (east) velocity, and circles are velocity averages determined by Pillsbury.

using all the GEK observations lying in that zone. These averages are tabulated in Table II, and plotted in Figure 3.

The mean downstream velocity profile is asymmetrical, with a cyclonic shear region about 15 nautical miles (28 km) wide having a shear of approximately 3×10^{-5} sec^{-1}. The anticyclonic shear region is about 32 nautical miles (59 km) wide and has a shear of approximately 2×10^{-5} sec^{-1}. For comparison, the value of the Coriolis parameter at this latitude is 6.9×10^{-5} sec^{-1}. The cyclonic shear of the average stream is significantly less than that of the instantaneous stream, because variations in position of the latter produce a wide distribution of average velocity, and hence a profile more gentle than that which would be found on any particular crossing.

3. Energy Calculations

The transfer of kinetic energy from the eddies to the mean flow can be expressed as

$$\varrho \overline{u'v'} \frac{\partial \overline{v}}{\partial x} \qquad (1)$$

where a bar represents a time average of a quantity, a prime represents a deviation from the time average, u and v are the velocity components in the cross-stream (x) and downstream (y) direction, and ϱ is the density of the water. This method for treating perturbations of a mean flow was first developed by O. Reynolds (1895), and is outlined by Lamb (1932).

Table 1

Onslow Bay					$k=1.46$				
Zone	$\bar{u}$ cm/sec		$\bar{v}$ cm/sec		$\overline{u'v'}$ cm²/sec²		$\frac{\partial \bar{v}}{\partial x}$ × 10-5sec-1	$\varrho\overline{u'v'}\frac{\partial \bar{v}}{\partial x}$ 10-2 ergs/cm³/sec	
6	+ 5.4	± 6.3	39.6	± 7.9	+ 210.2	± 279	2.13	+ 0.45	± 0.6
7	+ 2.6	± 7.2	53.6	± 13.0	— 72.1	± 503	2.68	— 0.19	± 1.3
8	+ 10.4	± 5.5	70.2	± 12.7	+ 236.3	± 341	4.35	+ 1.03	± 1.5
9	+ 6.9	± 6.6	95.9	± 9.1	+ 482.0	± 481	4.35	+ 2.10	± 2.1
10	— 4.8	± 8.0	118.1	± 8.8	+ 640.6	± 558	4.35	+ 2.79	± 2.4
11	+ 8.0	± 9.8	143.5	± 7.4	+ 648.8	± 364	3.84	+ 2.49	± 1.4
12	— 11.2	± 10.1	156.8	± 6.1	+ 335.6	± 409	0.71	+ 0.24	± 0.3
13	+ 2.2	± 6.0	157.2	± 6.4	+ 268.5	± 271	— 0.36	— 0.10	± 0.1
14	— 12.8	± 7.2	153.4	± 7.9	— 74.3	± 349	— 1.10	+ 0.08	± 0.4
15	+ 6.9	± 5.8	139.3	± 8.9	— 169.4	± 279	— 3.00	+ 0.51	± 0.8
16	— 3.8	± 5.4	126.0	± 8.3	— 61.9	± 253	— 2.21	+ 0.14	± 0.6
17	— 11.7	± 6.9	114.9	± 11.1	+ 72.4	± 349	— 2.85	— 0.21	± 1.0

Average Energy Flux: 79×10^{-4} ergs/cm³/sec.

Onslow Bay. For the Onslow Bay region, the data were divided into twelve zones across the current, each zone being 3 nautical miles wide. The eddy momentum transfer, $\overline{u'v'}$ was calculated by applying Simpson's rule to the value of $u'v'$, over the total time of observation, for each of the twelve zones. The value of $\frac{\partial \bar{v}}{\partial x}$ was calculated for each zone from the profile of average velocity ($\bar{v}$). Table 1 shows for each zone: $\overline{u'v'}$ the average transport of eddy momentum; $\frac{\partial \bar{v}}{\partial x}$ the shear of the average velocity; and the term (1) representing the production of mean kinetic energy by meanders. (ϱ was assumed constant, and equal to one gram per cubic centimeter.)

Figure 4 shows the distribution of (1) across the width of the current off Onslow Bay. The cross-stream integral of (1) is positive, indicating a net transfer of kinetic energy from the meanders to the mean current.

Florida Straits. The averages for the Florida Straits section were calculated by taking ensemble averages in each of the zones. That is, a barred quantity was evaluated as:

$$\overline{(\quad)} = \frac{1}{N}\sum_{i=1}^{N}(\quad)_i$$

where N is the total number of observations in the zone. The results of the calculations are tabulated in Table 2, and the profile of the term (1) across the current is shown in Figure 5. In this region, although there are points in the current where (1) is positive, the net production of mean kinetic energy across the whole stream is not significantly non-zero.

Statistical significance of results. The standard error of the means, $\bar{u}$, $\bar{v}$, and $\overline{u'v'}$ are given for each value in Tables 1 and 2. The standard error of a mean is defined, for large N, as $\sigma/\sqrt{N}$, where σ is the standard deviation of the sample from which the mean is calculated, and N is the number of individual observations.

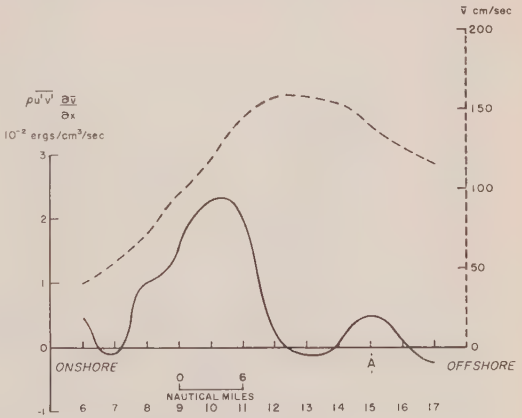


Fig. 4. Onslow Bay. Solid line: production of mean kinetic energy. Dotted line: mean downstream velocity.

Table 2

Florida Straits					$k=1.68$				
Zone	u cm/sec		$\bar{v}$ cm/sec		$\overline{u'v'}$ cm²/sec²		$\frac{\partial \bar{v}}{\partial x}$ $\times 10^{-5}$ sec^{-1}	$\varrho\overline{u'v'}\frac{\partial \bar{v}}{\partial x}$ 10^{-2}ergs/cm³/sec	
1	— 18.1	± 3.5	103.1	± 8.9	— 411.7	± 184	+ 4.27	— 1.76	± 0.78
2	— 17.6	± 6.4	129.9	± 10.4	+ 259.0	± 371	+ 5.68	+ 1.47	± 2.11
3	— 11.2	± 5.7	160.7	± 12.3	+ 270.0	± 392	+ 3.60	+ 0.97	± 1.41
4	— 6.4	± 4.5	164.8	± 5.8	+ 975.4	± 277	+ 4.27	+ 4.16	± 1.18
5	+ 11.0	± 6.7	202.8	± 9.2	+ 357.0	± 498	+ 1.97	+ 0.70	± 0.98
6	+ 9.9	± 4.2	188.0	± 4.5	+ 333.2	± 322	— 2.35	— 0.78	± 0.76
7	+ 16.5	± 5.7	167.0	± 7.4	+ 15.2	± 361	— 3.34	— 0.05	± 1.21
8	+ 14.4	± 4.2	131.4	± 5.5	— 13.2	± 217	— 3.55	+ 0.05	± 0.77
9	+ 5.1	± 4.4	107.2	± 6.6	+ 140.6	± 171	— 3.43	— 0.48	± 0.59
10	+ 5.8	± 4.5	73.5	± 6.9	+ 122.7	± 123	— 3.75	— 0.46	± 0.46
11	+ 2.1	± 4.4	43.9	± 7.4	+ 281.2	± 102	— 3.49	— 0.98	± 0.36

Average Energy Flux: 3.0×10^{-4} ergs/cm³/sec.

4. Discussion of results

Scale of perturbations. The conclusions which can be derived directly from the evaluation of (1) across the Florida Current apply only to the surface layer of moving fluid, and only to the time scale of perturbations which the u' and v' characterize. That is, the results obtained say nothing directly about the energy balance below the surface nor about the eddy motions of other time scales.

Pillsbury (1891) made a detailed study of the Florida Current at six anchor stations between Fowey Rocks and Gun Cay. His average surface velocities are plotted as circles in Figure 3 where they can be compared with the averages from the University of Miami

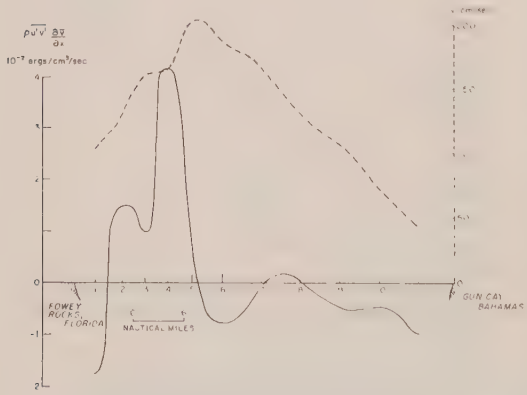


Fig. 5. Florida Straits. Solid line: production of mean kinetic energy. Dotted line: mean downstream velocity.

GEK measurements. His observations show that the subsurface velocities are not greatly less than those at the surface, and suggest a qualitative similarity between surface and subsurface velocity fields. If so, then the energy transfer between eddies and mean flow calculated for the surface layers from GEK data may be representative of most of the current depth in the Florida Straits.

Pillsbury also made a study of the time variations of current position, velocity, and width in the FloridaStraits, and found monthly variations related to the declination of the moon, as well as daily, tidally-influenced oscillations. He concluded that the tidal components of the variations were significantly larger than those which could be attributed to non-tidal causes.

Parr (1937) analyzed a set of five anchor stations which were successively occupied for 24-hour periods in the Straits of Florida between Fowey Rocks and Gun Cay. He found that lunar periodicities were strongly indicated in the data, and he even combined data from station to station, by referring them to corresponding lunar hours.

More recently, Murray (1952) analyzed velocity fluctuations in the same region. These fluctuations, as determined with the GEK, had periods between a few hours and a day. Murray was unable to confirm Pillsbury's conclusion that there was a relationship between the transport and the declination of the moon,

and found only inconclusive evidence of lunar effects. A further analysis of additional GEK data by HELA and WAGNER (1954) indicated that some tidal variations in velocity did exist, but that they were strongly masked by non-tidal effects. WERTHEIM (1954) from electromagnetic measurements of total transport, found conclusive evidence of diurnal tidal influence in the transport through the Florida Straits.

In conclusion, although it is probably rash to ascribe the velocity fluctuations of the current through the Straits of Florida predominantly to tidal causes, the periods of the fluctuations are of the order of a day, and the deviations from the mean which are used here to calculate the eddy transport of momentum most likely are representative of meandering motion.

In Onslow Bay, the data were taken continously for a month, in contrast to the Florida Straits, where the data were obtained at intervals over the course of several years. The eddy momentum transfer calculated from the perturbations has periods ranging between a few hours and a week; that is the range of meander periods which were observed (WEBSTER, 1961). Unfortunately, no systematic velocity measurements at depth, similar to those of Pillsbury in the Florida Straits, have ever been made in the Onslow Bay region.

We may conclude that both in the Florida Straits and in Onslow Bay, the calculated lateral surface transfer of momentum is produced by perturbations having periods of a day or longer. No observations have been made of perturbations of other time scales, in particular of small-scale perturbations, having periods which are small compared with the length of a day. Consequently, there is no evidence to suggest that small-scale perturbations would transport momentum in a manner similar to the meanders.

Regeneration Time. The magnitude and the lateral scale of the eddy kinetic energy release off Onslow Bay are similar to those in the Florida Straits. This apparent similarity between the two profiles suggests that there might be a characteristic scale of energy release throughout the whole region between the Straits of Florida and Cape Hatteras. However, a calculation of the time scales of the energy transfer reveals that the kinetic energy plays a different role in the mean kinetic energy balance in each region.

If the curves are integrated over their lengths, the average rate of surface transfer of energy from the meanders to mean flow is 79×10^{-4} ergs/cm^3/sec for Onslow Bay and 3.0×10^{-4} ergs/cm^3/sec for the Florida Straits. If the kinetic energy of the mean surface flow is averaged by integrating across the stream, the result is 7.3×10^3 ergs/cm^3 for Onslow Bay and 8.55×10^3 for the Florida Starits. If no other actions were present, the calculated rate of energy transfer would double the mean surface kinetic energy in 11 days in Onslow Bay and in 329 days in the Florida Straits.

The difference between the regeneration time of the surface kinetic energy in Onslow Bay and that in the Florida Straits may perhaps be explained partly by the confining channel of the latter. Near each shore there is a region where the meanders draw energy from the mean flow. This boundary layer effect is not so noticeable in the Onslow Bay region where the current is not closely confined by physical barriers.

In addition, the current through the Florida Straits may be more directly driven by a downstream pressure gradient. If so, then the meanders would contribute less to the mean flow than they do farther downstream where, possibly, the current is maintained by the cross-stream density field through the mechanism of meanders. At present, the measurements necessary to evaluate the role of downstream pressure gradients are not available.

Atmospheric Similarities. There is a similarity between the role of meanders in the Gulf Stream and the role of large-scale eddies in the atmosphere. Figure 6, adapted from STARR (1953) shows the production of zonal kinetic energy in arbitrary units, and the relative angular velocity as a function of latitude. When Figure 6 is compared with Figures 4 and 5, it can be seen that, for both systems, the maximum countergradient flow of momentum occurs in the region of maximum shear.

Several years' study of atmospheric transfer processes have established that the mean zonal flow is sustained by large-scale disturbances (STARR, 1953), and that the necessary meridional transports of momentum are effected by

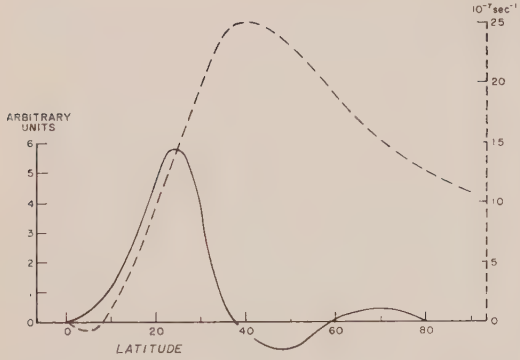


Fig. 6. Solid line: production of zonal kinetic energy, atmosphere. Dotted line: relative angular velocity. (from Starr, 1953, Figure 1.)

horizontal eddy exchange processes (Starr, 1954). The question of whether mean current systems in the ocean are maintained in an analogous manner cannot be answered until further observations are made.

Frictional models of the Gulf Stream (Stommel, 1948; Munk, 1950) require some sort of eddy dissipation which could conceivably be supplied by perturbations of meander scale. A coefficient of lateral eddy viscosity, or Austausch coefficient, may be defined on the assumption that the perturbations of the mean flow are analogous to an eddy frictional mechanism. However, the perturbations of the Gulf Stream flow, observed here, act exactly opposite to frictional effects, so that a viscosity coefficient calculated using them would be negative, and have questionable physical significance. It is possible that the necessary frictional dissipation is carried out by perturbations of a scale smaller than the meanders. If so, then the energy balance is analogous to that in the atmosphere, where the mean zonal flow is sustained by large-scale eddies, but dissipated by small-scale eddies and molecular viscosity.

5. Kinetic Energy Equation for the Mean Flow

The equation expressing the flux of energy in a turbulent flow was originally developed by O. Reynolds for a parallel non-rotating dissipative flow (Lamb, 1932). The corresponding general equation for the atmosphere has been given by Kuo (1951). The equations for ocean current systems are similar to those for the atmosphere because the coordinate frame for both systems is rotating, but dissimilar because the atmosphere is cyclic and centered about the axis of rotation, whereas ocean currents are a more local phenomenon. The action of winds upon the surface of the ocean is a mechanism whose atmospheric analogue is negligible.

The following model is a simple version of the Reynolds model, intended to provide an orientation. Better models may follow later.

Let us consider an ocean current, with a mean flow in the y-direction, having velocity components u, v, and w in the x, y, and z directions, x being directed to the right of y, and z being directed downward.

The momentum equations are

$$\varrho \frac{\partial u}{\partial t} + \varrho \mathbf{v} \cdot \nabla u = \varrho f v - \frac{\partial p}{\partial x} - X \qquad (1)$$

$$\varrho \frac{\partial v}{\partial t} + \varrho \mathbf{v} \cdot \nabla v = -\varrho f u - \frac{\partial p}{\partial y} - Y \qquad (2)$$

where ϱ is the density of the water, f is the Coriolis parameter, p is the pressure, and X and Y are components of a dissipative force. $\mathbf{v}$ is the three-dimensional velocity vector, $u\mathbf{i} + v\mathbf{j} + w\,\mathbf{k}$, and ∇ is the three-dimensional vector operator, $\mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z}$. We shall omit explicit reference to the wind stress on the surface. With the aid of the continuity equation, $\frac{\partial \varrho}{\partial t} + \nabla \cdot \varrho \mathbf{v} = 0$, we may write (1) and (2) as:

$$\frac{\partial \varrho u}{\partial t} + \nabla \cdot \varrho u \mathbf{v} = \varrho f v - \frac{\partial p}{\partial x} - X \qquad (3)$$

$$\frac{\partial \varrho v}{\partial t} + \nabla \cdot \varrho v \mathbf{v} = -\varrho f u - \frac{\partial p}{\partial y} - Y \qquad (4)$$

The quantities u, v, and w can be separated into mean motion and deviations as:

$$u = \bar{u} + u'$$
$$v = \bar{v} + v'$$
$$w = \bar{w} = w'$$

where

$$\overline{(\quad)} = \frac{1}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}} (\quad)\, dt.$$

The primes denote deviations from the time average. Terms of the form $\overline{u}\,\overline{v}$ become $u\,\overline{v} + u'v'$.

Let us now multiply (3) by $\overline{u}$ and (4) by $\overline{v}$, and take time averages:

$$\overline{u}\frac{\partial}{\partial t}\varrho\overline{u} + \overline{u}\nabla\cdot\varrho\overline{u}\,\overline{\mathbf{v}} + \overline{u}\nabla\cdot\varrho\overline{u'\mathbf{v}'} =$$

$$= \varrho f u\,\overline{v} - \overline{u}\frac{\partial\overline{p}}{\partial x} - \overline{u}\,\overline{X} \qquad (5)$$

$$\overline{v}\frac{\partial}{\partial t}\varrho\overline{v} + \overline{v}\nabla\cdot\varrho v\,\overline{\mathbf{v}} + v\nabla\cdot\varrho\overline{v'\mathbf{v}'} =$$

$$= -\varrho f\overline{u}\,\overline{v} - \overline{v}\frac{\partial\overline{p}}{\partial y} - \overline{v}\,\overline{Y} \qquad (6)$$

Equations (5) and (6) may be added, and with the aid of the continuity equation, written:

$$\varrho\frac{\partial}{\partial t}\overline{k} = -\,\overline{\mathbf{v}}\cdot\nabla\overline{k} - \overline{u}\nabla\cdot\varrho\overline{u'\mathbf{v}'} - \overline{v}\nabla\cdot\varrho\overline{v'\mathbf{v}'}$$

$$-\,\overline{u}\frac{\partial p}{\partial x} - \overline{v}\frac{\partial p}{\partial y} - d \qquad (7)$$

where d is the dissipative force, $\overline{u}\,\overline{X} + \overline{v}\,\overline{Y}$, and $\overline{k} = \dfrac{\overline{u}^2 + \overline{v}^2}{2}$ is the total kinetic energy of the horizontal mean current. By definition, $\dfrac{\partial}{\partial t}\overline{k} = 0$.

If (7) is integrated over a cross-section of the current in the x-z plane, the result gives the balance of horizontal kinetic energy of the mean flow in a unit cross-section of the current:

$$0 = \overset{①}{-\int\int}\overline{\mathbf{v}}\cdot\nabla\,\overline{k}\,dx\,dz$$

$$\overset{②}{-\int\int}\overline{u}\nabla\cdot\varrho\overline{u'\mathbf{v}'}\,dx\,dz$$

$$\overset{③}{-\int\int}\overline{v}\nabla\cdot\varrho\overline{v'\mathbf{v}'}\,dx\,dz$$

$$\overset{④}{-\int\int}\left(\overline{u}\frac{\partial\overline{p}}{\partial x} + \overline{v}\frac{\partial\overline{p}}{\partial y}\right)dx\,dz$$

$$\overset{⑤}{-\int\int} d\,dx\,dz \qquad (8)$$

With the limited amount of information available now, it is not possible to evaluate each term in equation (8) for the Gulf Stream. However, an estimate of the role of some of the terms can be made.

First, consider terms ② and ③:

$$\overline{u}\nabla\cdot\varrho\overline{u'\mathbf{v}'} + \overline{v}\nabla\cdot\varrho\overline{v'\mathbf{v}'} = \overline{u}\frac{\partial}{\partial x}\varrho\overline{u'u'} + \overline{u}\frac{\partial}{\partial y}\varrho\overline{u'v'}$$

$$+\,\overline{v}\frac{\partial}{\partial x}\varrho\overline{u'v'} + \overline{v}\frac{\partial}{\partial y}\varrho\overline{v'v'}$$

$$+\,\overline{u}\frac{\partial}{\partial z}\varrho\overline{u'w'} + \overline{v}\frac{\partial}{\partial z}\varrho\overline{v'w'}$$

The axes were chosen so that the mean current is in the y-direction, so that $\overline{v} \gg \overline{u}$. In addition, the perturbations travel in the y-direction, so that for the time averaged velocity perturbation terms, $\dfrac{\partial}{\partial y} \approx 0$. Hence, terms ② and ③ become:

$$u\nabla\cdot\varrho\overline{u'\mathbf{v}'} + \overline{v}\nabla\cdot\varrho\overline{v'\mathbf{v}'} \cong \overline{u}\frac{\partial}{\partial x}\varrho\overline{u'u'} + \overline{v}\frac{\partial}{\partial z}\varrho\overline{u'v'}$$

$$+\,\overline{u}\frac{\partial}{\partial z}\varrho\overline{u'w'} + \overline{v}\frac{\partial}{\partial z}\varrho\overline{v'w'}$$

The terms in $\overline{u'w'}$ and $\overline{v'w'}$ represent contributions of kinetic energy to the mean flow by vertical perturbations. Since observations of vertical motions are not available, this term has not been calculated. It might represent an important contribution to the mean down stream kinetic energy.

The term $\overline{v}\dfrac{\partial}{\partial x}\varrho\overline{u'v'}$ includes the expression used in the surface calculation of this study. It may be re-written as

$$\underset{\text{(a)}}{\frac{\partial}{\partial x}(\varrho\overline{v}\cdot\overline{u'v'})} - \underset{\text{(b)}}{\varrho\overline{u'v'}\frac{\partial\overline{v}}{\partial x}}$$

The term $\frac{\partial}{\partial x}\ (\varrho\bar{v}\cdot\overline{v'u'})$ represents an eddy advection across the boundaries of mean kinetic energy and can be integrated easily across a stream to become:

$$[\varrho\bar{v}\ \overline{v'u'}]_A - [\varrho\bar{v}\ \overline{v'u'}]_B$$

If the stream is bounded by walls, u' will be zero at the walls, or if the mean stream velocity drops to zero at each side of the current, $\bar{v}$ will be zero, and the term (a) is zero. Off Onslow Bay, where the current is not bounded by walls, and the observations did not cover the whole width of the current, calculation of the term (a) has revealed that its integral across the width of observations is zero. Hence the cross-stream integral of the term:

$$\varrho\overline{u'v'}\frac{\partial v}{\partial x} \qquad (1)$$

is a measure of the increase of mean kinetic energy at the expense of kinetic energy of horizontal eddy motion.

Maximum values of $\frac{\partial}{\partial x}\ \varrho\overline{u'u'}$ at the surface were observed to be almost as large as average values of $\frac{\partial}{\partial x}\ \varrho\overline{u'v'}$. However, since $v \gg u$, the term $u\ \frac{\partial}{\partial x}\ \varrho\overline{u'u'}$ is negligible in comparison with $\bar{v}\ \frac{\partial}{\partial x}\ \varrho\overline{u'v'}$.

Term ① represents the advection across the boundaries of the region under consideration of mean horizontal kinetic energy. Equation (8) states that this boundary advection ① must be balanced by the generation of mean kinetic energy by meanders, ② and ③, pressure forces, ④, and by the frictional dissipation, ⑤.

Proper measurements do not exist which will permit a calculation of the contribution of the pressure forces, ④, to the mean kinetic energy balance. Indications are that the downstream pressure gradient is important in the Florida Straits, but little is known about it in Onslow Bay. More reliable measurements are needed in both areas.

The frictional dissipation term, ⑤, was included in the equation to represent the general action of frictional stresses. Included in the term is the effect of wind stress at the surface. An order of magnitude calculation shows that the energy provided by the wind stress is one or two orders of magnitude less than the transfer of kinetic energy from the meanders to the mean flow. That is, if we assume a maximum wind stress of half a dyne/cm^2, a maximum velocity of 50 cm/sec, and a current depth of 500 meters, the energy transfer will be 5×10^{-4} ergs/cm^3/sec, for total correlation between winds and currents. Assuming only partial correlation, or more characteristic values of wind stress and meander velocity, the energy contributed to the meanders by the wind is about 1×10^{-4} ergs/cm^3/sec. In comparison with the transfer of energy between meanders and mean flow of 80×10^{-4} ergs/cm^3 sec, the wind would appear to be eliminated as a significant source of kinetic energy for the meanders.

6. Concluding Remarks

The similarity between these results and those in the atmosphere provides hope that a general study of the energy, momentum, and heat transfers in ocean currents would be as fruitful as such studies in the atmosphere have been. Perhaps, to some extent, meanders are to an ocean current what cyclones are to atmospheric circulation: a mechanism which sustains the mean flow and which provides transfers necessary for climatological balances.

A further program of systematic measurements in the Gulf Stream System is necessary to illuminate the large scale balances. Downstream from Cape Hatteras, especially, the tortuous course of the Gulf Stream has only recently been discovered, and the processes determining its nature are unknown. An extensive observation program might be undertaken initially with GEK surface velocity measurements. Hydrographic station data probably would not be suitable for velocity measurements because non-geostrophic and/or non-baroclinic velocity components might be an important constituent of the transfer process. A system which at present seems promising, and which ultimately might provide the vast

amount of data necessary for an oceanic study comparable to that in the atmosphere, is that of anchored buoys, which record data at predetermined time intervals. Such a system could be scattered throughout the ocean, and could record data at any depth.

Acknowledgements

The writer is indebted to Dr. Kirk Bryan, Jr. for suggesting this study. Thanks are also due to Prof. V. P. Starr for encouragement and helpful suggestions, and to Bruce Warren and Prof. William von Arx for helpful criticisms.

REFERENCES

Hela, I., and Wagner, L. P., 1954: *Note on tidal fluctuations in the Florida Current*, in Marine Lab., Univ. of Miami, Tech. Rept., No. 54—7, pp. 14—24.

Kuo, H. L., 1951: A note on the kinetic energy balance of the zonal wind systems. *Tellus*, **3**, pp. 205—207.

Lamb, Sir H., 1932: *Hydrodynamics*, 6th ed., Cambridge, England, Cambridge University Press, 738 pp.

Longuet-Higgins, M. S., Stern, M. E., and Stommel, H. M., 1954: The electrical field induced by ocean currents and waves, with applications to the method of towed electrodes. *Papers in Physical Oceanography and Meteorology*, **13**: 1, pp. 1—37.

Munk, W. H., 1950: On the wind-driven ocean circulation. *Jour. Meteor.*, **7**, pp. 79—93.

Murray, K. M., 1952: Short period fluctuation of the Florida Current from geomagnetic electrokinetograph observations. *Bull. Mar. Sci., Gulf and Caribbean*, **2**, pp. 360—375.

Parr, A. E., 1937: Report of hydrographic observations at a series of anchor stations across the Straits of Florida *Bull. Bingham Oceanogr. Coll.*, **6**, pp. 1—62.

Pillsbury, J. E., 1891: The Gulf Stream, a description of the methods employed in the investigation and the results of the research. *Rept. Supt., U. S. Coast and Geodetic Survey*, **1890**, App. 10.

Reynolds, O., 1895: On the dynamical theory of incompressible viscous fluids and the determination of the criterion. *Phil. Trans. Roy. Soc. London, A*, **186**, pp. 123—164.

Rossby, C.—G., 1936: Dynamics of steady ocean currents in the light of experimental fluids mechanics. *Papers in Physical Oceanography and Meteorology*, **5**:1, pp. 1—43.

Starr, V. P., 1953: Note concerning the nature of the large-scale eddies in the atmosphere. *Tellus*, **5**, pp. 494—498.

Starr, V. P., 1954: Commentaries concerning research on the general circulation. *Tellus*, **6**, pp. 268—272.

Starr, V. P., White, R. M., 1951: A hemispherical study of the atmospheric angular momentum balance *Quart. J. Roy. Meteor. Soc.*, **77**, pp. 215—225.

Stommel, H. M., 1948: The westward intensification of wind-driven ocean currents. *Trans. Amer. Geophys. Union*, **29**, pp. 202—206.

Stommel, H. M., 1958: *The Gulf Stream*. Berkeley, U.S.A., Univ. of California Press, 202 pp.

University of Miami, Marine Lab., 1954: *Some results of the Florida Current Study*, 1953. Technical Report No. 54—7, 100 pp.

University of Miami, Marine Lab., 1955: *Some results of the oceanographic studies in the Straits of Florida and adjacent waters*. Technical Report, No. 55—1, 74+ pp.

University of Miami, Marine Lab., 1957: *Tropical oceanography*, semi-annual report, March, 1957. Technical Report No. 57—9, 128 pp.

University of Miami, Marine Lab., 1958: *A technical report of investigations for the period of June 25, 1956 to May 7, 1957*. Technical Report, No. 58—3, 55 pp.

von Arx, W. S., 1950: An electromagnetic method for measuring the velocities of ocean currents from a ship under way. *Papers in Physical Oceanography and Meteorology*, **11**:3, pp. 1—62.

von Arx, W. S., 1954: Measurements of the oceanic circulation in temperate and tropical latitudes. *Oceanographic Instrumentation, Nat. Res. Counc.* (U.S.) Publ., No. **309**, pp. 13—35.

von Arx, W. S., Bumpus, D. F., and Richardson, W. S., 1955: On the fine structure of the Gulf Stream front. *Deep Sea Res.*, **3**, pp. 46—65.

Wertheim, G. K., 1954: Studies of the electric potential between Key West, Florida and Havana, Cuba. *Trans. Amer. Geophys. Union*, **35**, pp. 872—882.

Webster, T. F., 1961: A description of Gulf Stream meanders off Onslow Bay. *Deep-Sea Res.* (in press).

Note on the Propagation of Baroclinic Waves in a Zonal Ocean Current

By PIERRE WELANDER, Swedish Research Council of the Natural Sciences and International Meteorological Institute in Stockholm

(Manuscript received June 29, 1961)

Abstract

In a zonal ocean current one sees stationary or slowly moving baroclinic waves, with a wavelength of 1,000 km or more. These waves may be equivalent to a type of very long waves studied earlier in the atmospheric case. The wave speed is independent of the wave number but depends on the speed of the basic current, the stratification etc. For the case of a zonal current with exponential vertical profile, stationary waves are found for an eastward current. The wave perturbation normally decays downward in an exponential way. The parameter that decides the nature of the waves is $\lambda = \frac{\beta}{f^2}\frac{\Delta A\, p_0}{U_0}$ where U_0 is the amplitude of the basic current, ΔA the vertical variation of the specific volume, p_0 the scale pressure (the pressure variation that corresponds to the scale depth of the basic current), and f the Coriolis parameter. For a westward current travelling neutral waves exist but only for special values of λ that are not normally expected in the oceans.

1. Introduction

In the open ocean, far away from the coasts, we find a system of essentially zonal ocean currents. The vertical structure of the currents is relatively uniform, and as a normal form we may take a velocity distribution with maximum velocity at the surface and an exponential-type decay downwards. Also the stratification can normally be described as exponential-type. At great depth we find only small motions and the water is practically homogeneous.

It is of interest to investigate theoretically the existence of different waves in such a current. The present investigation deals only with a special type of wave, namely low-frequency baroclinic waves having wavelengths of 1,000 km or more. These waves are the ones mainly observed in synoptic studies of the ocean circulation. Theoretically such waves have been studied in a resting two-layer ocean by VERONIS and STOMMEL (1956).

The present theory is based on the equations for very long waves in the atmosphere used earlier by the author (WELANDER, 1961). The basic equations are practically identical for the atmosphere and the ocean if the so-called p-system is used. For the ocean we may replace the law of conservation of potential temperature by the simpler law of conservation of density or specific volume. The boundary conditions are that $\omega = 0$ at the free surface ($p = 0$), and that $\phi' \to 0$ at great depths ($p \to \infty$).

The effect of the bottom is not considered here, and we allow vertical velocities to exist at infinity. For application to a real ocean we should think of the model as a boundary layer one describing the conditions in the upper stratified part of the water. It is fitted to a deep water solution in which the vertical velocity decreases linearly with depth to a zero value at the bottom. Some justification for this procedure is found in a recent study of the thermocline by H. Stommel (private communication).

In the earlier mentioned paper a simplified wave equation was obtained for low-frequency waves of planetary scale (wave-lengths many thousand kilometers), when the Rossby number (Ro) and the inverse Richardson number (Ri^{-1}) characterizing the basic state were assumed small. In the atmosphere Ro would be of the order 10^{-2} to 10^{-1}, and Ri^{-1} of the order 10^{-2}. In the open ocean we find that Ro is of the order 10^{-4} to 10^{-3} and Ri^{-1} of the order 10^{-3}. The smallness of these coefficients in the oceanic case means that the same wave equation can be applied also to shorter wave-lengths, possibly down to a wave-length of 1,000 km.

2. General equations

Let x and y be horizontal coordinates related to the latitude ϕ, the longitude λ, and to the radius of the earth R as follows:[1]

$$\begin{aligned} dx &= R\cos\phi\, d\lambda \\ dy &= R\, d\phi \end{aligned} \tag{1}$$

Further let p be the pressure, α the specific volume $\left(\alpha = \frac{1}{\varrho}\right)$, $\phi = gz$ the geopotential of a pressure surface, and let u, v, ω be the velocity components in the x-y-p system $\left(\text{thus } \omega = \frac{dp}{dt}\right)$. Assuming that the water is ideal and incompressible and that a hydrostatic balance exists in the vertical direction, and assuming further that the Rossby number is so small that we may neglect spherical acceleration terms of the type $\frac{\operatorname{tg}\phi}{R} uv$ etc., the equations governing u, v, ω, α and ϕ are

$$\frac{du}{dt} - fv = -\frac{\partial \phi}{\partial x} \tag{2}$$

$$\frac{dv}{dt} + fu = -\frac{\partial \phi}{\partial y} \tag{3}$$

$$\alpha = -\frac{\partial \phi}{\partial p} \tag{4}$$

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} - \frac{\operatorname{tg}\phi}{R} v + \frac{\partial \omega}{\partial p} = 0 \tag{5}$$

$$\frac{d\alpha}{dt} = 0 \tag{6}$$

where $\frac{d}{dt} = \frac{\partial}{\partial t} + u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y} + \omega\frac{\partial}{\partial p}$.

The basic current U is assumed to be uniform in the y-direction. The basic fields of the geopotential Φ and the specific volume A satisfy the relations

$$fU = -\Phi_y \tag{7}$$

$$A = -\Phi_p \tag{8}$$

(in the following derivatives of the basic quantities are denoted by a subscript).

We write

$$\left.\begin{aligned} u &= U + u' \\ v &= v' \\ \omega &= \omega' \\ \alpha &= A + \alpha' \\ \phi &= \Phi + \phi' \end{aligned}\right\} \tag{9}$$

and derive the perturbation equations. Using an expansion of the same type as in the atmospheric case, but with the Rossby number and inverse Richardson number defined as

$$\left.\begin{aligned} Ro &= \frac{U_0}{Rf} \\ Ri^{-1} &= \frac{U_0^2}{p_0 \Delta A} \end{aligned}\right\} \tag{10}$$

where U_0 is the amplitude of the basic current, p_0 the scale depth, and ΔA the vertical variation of the specific volume, a correspondingly simplified wave equation is obtained. The

[1] Although a β-plane analysis will give results that are correct to the first order, the quasi-coordinates x-y-p that depict the spherical geometry correctly have been used throughout for the sake of consistency in this paper. Thus the term $\frac{\operatorname{tg}\phi}{R} v$ is kept in the continuity equation, and the permutation rule

$$\frac{\partial^2}{\partial y \partial x} = \frac{\partial^2}{\partial x \partial y} + \frac{\operatorname{tg}\phi}{R}\frac{\partial}{\partial x}$$

that is valid for the quasi-coordinates has been applied.

values of Ro and Ri^{-1} quoted earlier will be obtained by choosing $U_0 \sim 10$ cm sec^{-1}, $p_0 \sim 10^8$ dyn cm^{-2} (scale depth 1,000 m), $\Delta A \sim 10^{-3}$ g^{-1} cm^3, $f \sim 10^{-4}$ sec^{-1}. The simplified wave equation is equivalent to the system of perturbation equations in which the two equations of motion are replaced by the geostrophic condition for the perturbation:

$$-fv' = -\frac{\partial \phi'}{\partial x} \tag{11}$$

$$fu' = -\frac{\partial \phi'}{\partial y} \tag{12}$$

$$\alpha' = -\frac{\partial \phi'}{\partial p} \tag{13}$$

$$\frac{\partial u'}{\partial x} + \frac{\partial v'}{\partial y} - \frac{\operatorname{tg}\phi}{R} v' + \frac{\partial \omega'}{\partial p} = 0 \tag{14}$$

$$\frac{\partial \alpha'}{\partial t} + U\frac{\partial \alpha'}{\partial x} + fU_p v' + A_p \omega' = 0 \quad (A_y = -\Phi_{yp} = fU_p) \tag{15}$$

Solving ω' from (15) and using (11) and (13) we get

$$\omega' = \frac{1}{A_p}\left\{\frac{\partial^2 \phi'}{\partial p \partial t} + U\frac{\partial^2 \phi'}{\partial x \partial p} - U_p \frac{\partial \phi'}{\partial x}\right\} \tag{16}$$

and inserting the velocity values in the continuity equation yields

$$\frac{\partial}{\partial p}\left\{\frac{1}{A_p}\left(\frac{\partial^2 \phi'}{\partial p \partial t} + U\frac{\partial^2 \phi'}{\partial x \partial p} - U_p \frac{\partial \phi'}{\partial x}\right)\right\} - \frac{\beta}{f^2}\frac{\partial \phi'}{\partial x} = 0. \tag{17}$$

The boundary conditions, requiring that ω' vanish at $p = 0$ and that ϕ' vanish for infinite pressures, become

$$\left.\begin{aligned} \frac{\partial^2 \phi'}{\partial p \partial t} + U\frac{\partial^2 \phi'}{\partial x \partial p} - U_p \frac{\partial \phi'}{\partial x} = 0 \quad & \text{at } p = 0 \\ \phi' = 0 \quad & \text{at } p = \infty \end{aligned}\right\} \tag{18}$$

3. The wave form

Introducing a wave form

$$\phi' = \hat{\phi}(y, p)\, e^{ik(x - ct)} \tag{19}$$

where k is the wave number and c the wave velocity (positive for eastward motion) equations (17) and (18) transform into

$$(U - c)\frac{\partial}{\partial p}\left(\frac{1}{A_p}\frac{\partial \hat{\phi}}{\partial p}\right) - \left\{\left(\frac{U_p}{A_p}\right)_p + \frac{\beta}{f^2}\right\}\hat{\phi} = 0 \tag{20}$$

$$\left.\begin{aligned} (U - c)\frac{\partial \hat{\phi}}{\partial p} = U_p \hat{\phi} \quad & \text{at } p = 0 \\ \hat{\phi} = 0 \quad & \text{at } p = \infty \end{aligned}\right\} \tag{21}$$

(These equations can be directly identified with the equations (30) and (32) in the earlier paper, if we substitute A for A_p.)

The wave number k does not occur in the wave equation or the boundary conditions, hence the wave speed must be independent of k. Further we note that no y-derivatives appear, and this makes it possible to treat the y-variable parametrically.

We consider in particular the basic profiles (Fig. 1)

$$U = U_0 e^{-\frac{p}{p_0}} \tag{22}$$

$$A = A_0 + \Delta A e^{-\frac{p}{p_0}} \tag{23}$$

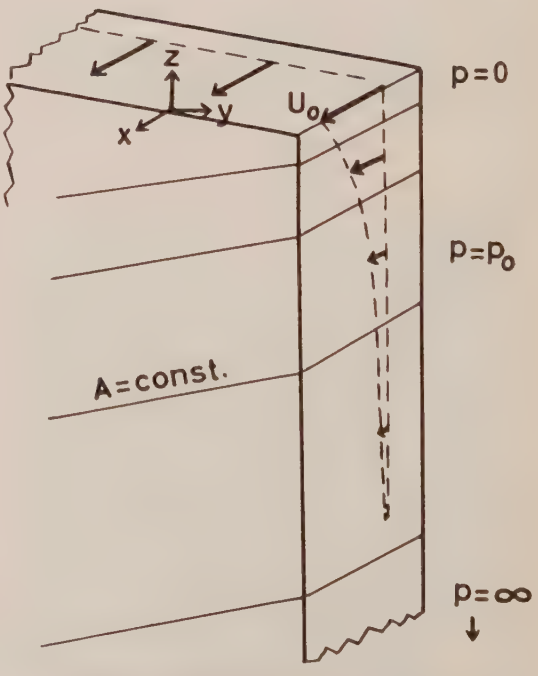


Fig. 1. Sketch of the basic current.

where p_0 is the scale pressure (this must necessarily be the same for U and A in view of the "thermal wind" equation). AA will show some variation in the y-direction, but we study here only a "local" solution, in which y is fixed. With the notation

$$\left.\begin{aligned} P &= \frac{p}{p_0} \\ c^\star &= \frac{c}{U_0} \\ \lambda &= \frac{\beta}{f^2}\frac{\Delta A\, p_0}{U_0} \end{aligned}\right\} \qquad (23)$$

the equations (20) and (21) change to

$$(e^{-P} - c^\star)\frac{d}{dP}\left(e^P \frac{d\hat{\phi}}{dP}\right) + \lambda\hat{\phi} = 0 \qquad (24)$$

$$\left.\begin{aligned} (1 - c^\star)\frac{d\hat{\phi}}{dP} + \hat{\phi} &= 0 \qquad \text{at } P = 0 \\ \hat{\phi} &= 0 \qquad \text{at } P = \infty \end{aligned}\right\} \qquad (25)$$

4. The stationary waves

In this case $c^\star = 0$ and (24) and (25) reduce to

$$e^{-P}\frac{d}{dP}\left(e^P \frac{d\hat{\phi}}{dP}\right) + \lambda\hat{\phi} = 0 \qquad (26)$$

$$\left.\begin{aligned} \frac{d\hat{\phi}}{dP} + \hat{\phi} &= 0 \qquad \text{at } P = 0 \\ \hat{\phi} &= 0 \qquad \text{at } P = \infty \end{aligned}\right\} \qquad (27)$$

Attempting a solution of the form $e^{\alpha P}$ gives $\alpha(\alpha+1) + \lambda = 0$, thus the general solution is of the form

$$\hat{\phi} = C_1 e^{\alpha_1 P} + C_2 e^{\alpha_2 P} \qquad (28)$$

$$\alpha_1 = -\frac{1}{2} + \sqrt{\frac{1}{4} - \lambda},\quad \alpha_2 = -\frac{1}{2} - \sqrt{\frac{1}{4} - \lambda} \qquad (29)$$

If $\lambda < 0$, as is the case when $U_0 < 0$, the second boundary condition requires that $C_1 = 0$, and the first condition gives then $\alpha_2 = -1$, i.e. $\lambda = 0$. No stationary waves occur. If $\lambda > 0$ the second condition is fulfilled and applying the first condition we find possible solutions of the form

$$\hat{\phi} = \text{const}\left\{(1+\alpha_2)\, e^{\alpha_1 P} - (1+\alpha_1)\, e^{\alpha_2 P}\right\} \qquad (30)$$

For, $\lambda < \frac{1}{4}$ the profiles are of exponential type, for $\lambda > \frac{1}{4}$ of oscillatory type. In the real ocean λ is generally small and we are likely to have solutions of the first type.

5. Travelling neutral waves

In the case $c \neq 0$ we introduce the new independent variable

$$\zeta = \frac{1}{c^\star} e^{-P} \qquad (31)$$

This gives

$$\zeta(\zeta - 1)\frac{d^2\hat{\phi}}{d\zeta^2} + \lambda\hat{\phi} = 0 \qquad (32)$$

$$\left.\begin{aligned} (1 - c^\star)\frac{d\hat{\phi}}{d\zeta} - c^\star\hat{\phi} &= 0 \quad \text{at } \zeta = \frac{1}{c^\star} \\ \hat{\phi} &= 0 \qquad \text{at } \zeta = 0 \end{aligned}\right\} \qquad (33)$$

The solution that is regular at $\zeta = 0$ is given by the series

$$\hat{\phi} = \zeta + \frac{1}{2}\lambda\zeta^2 + \frac{1}{2}\lambda\frac{\lambda + 1\cdot 2}{2\cdot 3}\zeta^3 + \\ + \frac{1}{2}\lambda\frac{\lambda + 1\cdot 2}{2\cdot 3}\frac{\lambda + 2\cdot 3}{3\cdot 4}\zeta^4 + \ldots \qquad (34)$$

The second boundary condition is automatically satisfied. The solution is not well behaved at $\zeta = 1$, except when $\lambda = -n(n+1)$, $n = 1, 2, 3, \ldots$, and the solution becomes a polynomial[1].

In other cases, we have thus to restrict ourselves to the region $\zeta < 1$, and it must be required that $c^\star > 1$, or $c > U_0$.

The travelling neutral waves can be found

[1] The general solution to the equation (32) is $|(2\zeta-1)^2-1|\, P_\nu'(2\zeta-1)$, where P_ν is the Legendre function of order ν, and $\lambda = -\frac{1}{4}\nu(\nu+1)$. *Cf.* Kamke, Differential-Gleichungen I, eq. No. 2.231.

by the following graphical construction. We draw the family of solutions $\hat{\phi}\ (\lambda, \zeta)$ given by (35) together with the curves $\hat{\phi}^{\star} = \text{const.}$ $(\zeta - 1)$ and seek for points at which the solution curves are tangent to the $\hat{\phi}^{\star}$-curves.

The abscissa of such a point represent $\frac{1}{c^{\star}}$. The construction is easily understood if we note that $\hat{\phi}^{\star}$ satisfies the relation $(\zeta - 1)\frac{d\hat{\phi}^{\star}}{d\zeta} - \hat{\phi}^{\star} = 0$ that for $\zeta = \frac{1}{c^{\star}}$ is precisely the first boundary condition (33).

Solutions are found for $\lambda \leq -2$, i.e. for a westward current. The waves travel westward with a speed of the order U_0. In the polynomial case one wave remains stationary relative to the surface current.

6. Application to the ocean

With the previously given values for U_0, p_0, and ΔA, and with f and β evaluated at middle latitudes we find a value of λ around 0.1. There may, of course, be reasons to choose other values. In the surface layer of the ocean we may have scale pressures that are one order of magnitude smaller, and λ will then be decreased correspondingly, or we may consider a strong current of 100 cm sec^{-1} that also would require a λ-value one order of magnitude smaller. Larger values of λ may be needed when the current is very weak, or when we come close to the equator (in this region the theory will, however, break down for other reasons). Under normal conditions we should expect the analysis for small λ-values to be well applicable. The general conclusions we could draw from the investigations are then: Stationary baroclinic waves can occur in an eastward current while no stationary or travelling waves are expected in a westward current of the present type.

REFERENCES

Veronis, G. and Stommel, H., 1956: The action of variable wind stresses on a stratified ocean. *Journal of Marine Research*, Vol. **15**, p. 43.

Welander, P., 1961: Theory of the Long Waves in a Zonal Atmospheric Flow. *Tellus* **13**, p. 140.

Measurements of the Water Vapour, Tritium and Carbon-14 Content of the Middle Stratosphere over Southern England

By F. BROWN[1], P. GOLDSMITH[2], H. F. GREEN[1], A. HOLT[1] and A. G. PARHAM[2]

(Manuscript received May 8, 1961)

Abstract

Measurements of the water vapour, tritium and carbon-14 content of the stratosphere at heights of between 80,000 and 100,000 feet, made over England during the years 1956 to 1960, are described.

The tritium and carbon-14 concentrations are greater than those expected from natural production due to the cosmic radiation. The bulk of these two isotopes, at present in the stratosphere, has been injected there during the course of thermonuclear explosions.

Mass spectrometric analyses show that the deuterium/hydrogen ratio of the water and the carbon-13/carbon-12 and oxygen-16/oxygen-18 ratios of the carbon dioxide collected at these heights were the same, within the accuracy of measurement, as those for samples collected near the ground.

The concentration of tritium at the heights of the measurements has not fallen since the termination of thermonuclear weapon testing in autumn 1958. There is poor correlation between the tritium concentration at 90,000 feet and that of particulate caesium-137 at 47,000 feet measured over the sampling period. The humidity above 80,000 feet is found to be greater than that in the lower stratosphere.

It is concluded from these humidity and tritium results that the large scale meridional circulation in the stratosphere, first postulated by Brewer and Dobson, does not reach heights of 90,000 feet over temperate latitudes.

1. Introduction

During the years 1956 to 1960, a series of balloon flights at altitudes of 80,000—100,000 feet were carried out over southern England. Samples of water vapour and carbon dioxide were collected by drawing stratospheric air through a vapour trap cooled in liquidnitrogen. A radio-command or clock-controlled system operated the equipment and subsequently released it from the balloon on a parachute.

The samples of water and carbon dioxide, so obtained, were separated, measured and then assayed by conventional low background counting techniques to give their tritium and carbon-14 contents respectively. Mass spectrometric analyses of the gases yielded data on the oxygen-18/oxygen-16, carbon-13/carbon-12 and deuterium/hydrogen ratios.

Since the carbon dioxide concentration of the atmosphere, remote from the ground, is constant (GLUECKAUF, 1944, HAGEMANN ET AL., 1959), the amount of carbon dioxide collected on a flight is used as a measure of the total quantity of air sampled. Thus, the tritium and carbon-14 concentrations of the stratospheric air are deduced, together with the humidity mixing ratio.

[1] Atomic Weapons Research Establishment, Aldermaston, England.

[2] Atomic Energy Research Establishment, Harwell, England.

Preliminary results on the 1956—1958 flights have already been reported (BARCLAY ET AL., 1960, GOLDSMITH ET AL., 1960). These, together with later, more extensive, results, are presented and discussed in this paper.

2. Experimental Techniques

2.1. *The Balloon Flights*

The sampler and associated equipment (basket, parachute, radar reflectors, radio transmitters, radio command receiver and helicopter homing equipment) weighed approximately 200 lbs and were carried on a polythene balloon of the type developed by the Cosmic Ray Research Group of the University of Bristol, England. These balloons were non-expansible, open necked and about 180 feet long. The rate of ascent was approximately 800 ft/min. and so ceiling (~ 90,000 feet) was reached in about 2.5 hours when the hydrogen gas had expanded to fill completely the polythene envelope. The balloon then floated and any further height changes were caused by gas leaks or fluctuations in the gas temperature. These changes were usually quite small over three to four hours and the sampling was carried out during most of this time. The load and parachute were released from the balloon either by a preset clockwork trigger or a radio command system operated from the ground.

All balloons were launched from Cardington, near Bedford, England. There were three essential meteorological conditions required on the day of a flight.

a) Light surface winds during launching operations, to facilitate balloon handling.
b) A vertical wind profile such that the balloon had not drifted over the sea by the end of the sampling period.
c) Clear skies so that the balloon did not rise through cloud and so introduce the possibility of tropospheric water being deposited on the equipment.

In general condition (b) can only be met if the stratospheric wind at the sampling height is easterly. This limits the period when flights of long duration can be made over England to the spring and summer months, and the number of days in this period which meet the other flight conditions are very limited.

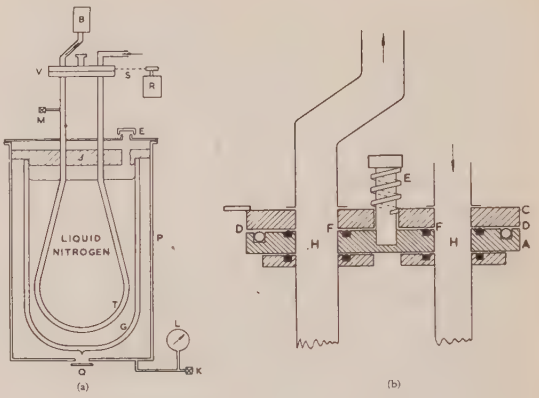


Fig. 1. (a) The vapour trap; (b) details of rotary valve V.

The balloons were tracked by Radar and recovery was by car, although on some occasions helicopters were used.

2.2 *Description and Operation of the Collection Apparatus*

The sampling apparatus, which has been fully described elsewhere (GOLDSMITH and PARHAM, 1960) is shown schematically in figure 1a. The essentials were a stainless-steel tubular vapour-trap T, a Dewar flask G containing liquid nitrogen, a centrifugal blower B and a disc-type rotary valve V which is detailed in figure 1b.

The air entered the sampler through the intake pipe connected to a port in the upper movable disc C of the rotary valve V. It was then drawn through the vapour-trap tube T, via the outlet port of the valve, by the blower B.

The valve V, consisted of two polished aluminium plates pressed together by a spring E and slightly separated by a ring of steel balls D located in a channel in the lower plate A. A vacuum seal was maintained by neoprene rubber-ring seals F, likewise located in channels in the lower plate A, one around each port H. The valve was then operated by rotating the upper plate C with respect to the lower fixed plate A by a rotary actuator R.

A wide-bore valve of this type was chosen to avoid constrictions, to minimise the use of rubber and grease, and to facilitate the baking-out of as much of the inner surfaces of the trap as possible.

Prior to the day of a flight, the vapour-trap was pumped out, through the all-metal valve M, with the valve V in the closed position, to a pressure of $\sim 10^{-5}$ mm Hg, whilst the tube was baked out in a special oven at a temperature of $\sim 400°$ C. The trap was then placed in the aluminium pressure vessel P which already contained the empty Dewar flask G. The non-conducting lid J of the flask was then carefully positioned and the stainless steel top of the pressure vessel, which was an integral part of the vapour-trap tube, bolted down. A rubber ring seal ensured that the vessel was gas-tight.

The vapour-trap was then pumped continuously until the day of the flight, when the trap was sealed at M. Only then was the liquid nitrogen loaded into the sampler through the filler plug E. Precautions were taken to avoid the collection of oxygen and to prevent the solidification of the liquid nitrogen when boiling under reduced pressure. This was achieved by maintaining the pressure inside P at about one atmosphere above that of the environment using the ball release valve K. K was adjusted before the flight, and so at full altitude, the absolute pressure inside P was about one atmosphere. This excess pressure was maintained by slow evaporation of the liquid nitrogen in the Dewar flask. A large-bore safety valve Q was also incorporated in case breakage of the Dewar flask, on launch or landing, resulted in rapid evaporation of the liquid nitrogen. As a further precaution, Q was opened permanently on closure of the valve V after sampling. As an additional safeguard Q could be tripped by a pressure capsule which was set to trigger at about 5,000 feet on the descent.

The sampler was powered by a nominal 24 volt silver-zinc accumulator, contained in a sealed housing. The opening and closing of the valve at altitude was determined by a preset clockwork mechanism or the radio command system.

The position of the valve, the performance of the blower and the pressure above the liquid nitrogen were all monitored by a modified radiosonde carried with the equipment. A Pressure-sonde, developed at the University of Bristol, was also flown and used to calculate the height of the balloon (an I.C.A.N. atmosphere was assumed).

2.3 *Precautions, Tests and Calibrations*

Every effort was made to minimize the amount of water carried up on the equipment. The external surfaces of the sampler were maintained at temperatures above the dew point of the surrounding air, by effective thermal insulation around the liquid nitrogen and because the upper parts of the sampler were black, so enhancing direct solar heating. These conditions did not apply on the first two flights in 1956 and 1957, which resulted in excessive water being collected.

Many other possible sources of spurious water were checked, the cane basket was replaced on later flights by a metal frame without affecting the measured humidity. Care was always taken that no components of the equipment were allowed to come into contact with grass of the launching field.

The most obvious sources of spurious water were the balloon and the hydrogen gas. Cylinders of dry gas were used to lessen the latter effect. The length of cord between sampler and balloon was varied between 200 and 500 feet without affecting the amount of water collected.

A more elegant check was made by measuring the deuterium/hydrogen ratio of the hydrogen of the collected water in 1959. On the 1960 flights several grams of heavy water (D_2O) were added to the balloon gas, this approximately doubled the amount of water in the gas and so made the deuterium/hydrogen ratio of this water about 2. The deuterium/hydrogen ratio of the water collected at altitude, however, remained unchanged at 2×10^{-4} which conclusively showed that diffusion of water from balloon to sampler was not important. These deuterium/hydrogen ratios are given with the other Mass Spectrometric results (Table 3).

The minimum length of the tube in the cold trap was calculated using the well established solutions to the analogous heat-conduction equations (Carslaw and Jæger, 1947). Uncertainty in the collection efficiency for molecules of carbon dioxide and water at the cold surface as well as the possibility of mist formation in the tube made it seem that a collection efficiency of less than 100 % for the two gases was possible. Measurements were therefore made of the flow rate through the sampler

Table 1

Flight Date	Height 10^3 feet	Volume of CO_2 ccs NTP	Sampling Time Mins.	Flow rate from CC_2 litres/sec	Calibrated Flow rate Litres/sec
13.5.59	91	95	160	1.6	2.3
5.6.59	80	150.1	110	2.2	2.7
15.6.59	93	69.0	51	3.8	2.3
18.6.59	90	157.9	150	2.9	2.4
7.7.60	89	139.6	125	2.7	2.4
11.9.60	99	76.5	133	2.2	< 2.3

in a large decompression chamber using a hot-wire anemometer. In Table 1, the results of these calibrations are compared with the flow rates on actual flights, as deduced from the amount of carbon dioxide collected, the duration and height of the flight and by assuming an air temperature of −53° C. The carbon dioxide concentration for this purpose is taken as 0.031 % by volume (HAGEMANN et al., 1959), not 0.025 % as found in earlier work (GLUECKAUF, 1944).

The flow rates found on the flight were more variable than those found on calibration. This is not unreasonable since: —

a) the load imposed on the blower motor by the monitoring reduction gear was probably variable.
b) different blowers of the same model were used on each flight.
c) the calibrated flows are appropriate to an accumulator voltage of 27, whilst in fact the potential applied to the blower during a flight was 30 volts falling to 24 volts after about one hour.

The mean flow rate on the first five of the above flights is 2.6 litres/sec, compared with a mean calibrated flow of 2.4 l/sec. This suggests that the vapour collection in the sampler was complete.

2.4 *Gas Analysis*

On recovery of the sampling equipment after a flight the contents which consisted of the collected carbon dioxide and water vapour together with air at the ambient pressure of the sampling altitude, were transferred to a storage tube to await detailed analysis.

The recovered sampler was connected to a vacuum apparatus and the gases drawn from the trap, via a needle valve, through two pyrex glass U-tubes connected in series and cooled in liquid nitrogen. The gas flow rate was adjusted so that the pressure of the non-condensable gases on the exit side of the U-tubes did not exceed 0.05 mm Hg. Under these conditions complete condensation of the carbon dioxide and water was achieved as verified by the fact that the second U-tube never contained any condensed vapours. Any vapour absorbed on the walls of the sampler tube was removed by raising the temperature to approximately 400° C and pumping until the pressure fell to the order of 10^{-5} mm Hg. The total transfer was completed in three hours, whereupon the U-tubes were sealed by melting constructions in the arms. Subsequent entry to the U-tubes was via break seals.

In the laboratory the U-tube was connected to a high vacuum apparatus and the break seal opened. The gas sample was passed slowly through a nitrogen cooled trap to remove any last traces of noncondensable gases.

Separation of the carbon dioxide from the water was achieved by repeated distillation from frozen pentane (MP$\sim$ −130° C, V. P. $H_2O < 10^{-7}$ mm Hg. V. P. $CO_2 \sim 1$ mm. Hg). The carbon dioxide was then pumped into a calibrated volume and samples taken in small break seal tubes for mass spectrometric analysis. The water was condensed into a tube containing approximately 10 g of zinc, which had been outgassed by heating at 400° C under vacuum for several hours. This tube was then sealed and heated at 400° C for 36 hours to effect complete reduction of the water. On completion of reduction the tube was opened, the hydrogen transferred into a calibrated volume and samples taken for mass spectrometric analysis. The amount of water

originally collected was calculated from the volume of hydrogen obtained and calibration runs with known amounts of water indicated that the inaccuracies would not be > 1 % using this procedure. Known amounts of the hydrogen and the carbon dioxide were pumped into proportional counter for tritium and carbon-14 determinations respectively.

2.5 *Counting*

Tritium

The tritium was measured in a proportional counter. The counter, 3 inches in diameter and 12 inches long, was constructed of copper with a 1 mil tungsten anode centre wire and had a sensitive volume of approximately 1 litre. To reduce the natural background and so increase the sensitivity of the apparatus, a double layer of Geiger counters was arranged round the counter and operated in anti-coincidence. The whole counting assembly was mounted in a steel castle with 8 inch thick walls and 8 inch thick sliding doors. The counter was followed by conventional amplifier anticoincidence and scaler units. Backgrounds of the order of 10 counts/minute were obtained and were at least an order of magnitude less than the activity of the measured samples.

The hydrogen samples were introduced into the counter and the pressure raised to 1 atmosphere with the counting gas (argon/10 % methane). Provided that the partial pressure of hydrogen was less than 10 cm. Hg. the resolution and plateau characteristics of the counter were unimpaired. Samples were counted to 1 % statistics. Counter end effects had been determined experimentally and the necessary corrections were applied. A Fermi plot of the β spectrum was consistent with all counts being due to tritium.

The 1956 sample was also counted independently by Geiger Counters (Gard, 1958), and satisfactory agreement was found.

Radiocarbon

The radiocarbon was determined in a similar manner but the counter was smaller having a sensitive volume of approximately 250 cc. Backgrounds of approximately 2 counts/minute were obtained. Since the carbon dioxide samples were small, the radiocarbon activity was less than the background. Under these conditions the accuracy was poor and standard deviations of ± 8 %— ± 60 % were the best that could be obtained.

2.6 *Mass Spectrometric Analyses*

Analyses of the carbon dioxide and hydrogen samples yielded data on the carbon-13/carbon-12, oxygen-18/oxygen-16 and deuterium/hydrogen ratios respectively (Table 3).

A Metro-Vic Mass Spectrometer Type MS_2 was used which had a 90° tube, 15 cm. radius ion path and an acceleration voltage of 2KV. It was fitted with a model 51A Vibron MS vibrating reed type amplifier to increase the sensitivity.

To correct for instrumental variations over the period, samples of the same standard carbon dioxide or hydrogen were measured before and after each stratospheric sample.

3. Results

3.1 *Tritium, radiocarbon and humidity*

The results of the tritium, radiocarbon and humidity determinations are summarised in Table 2. Basic measurements are given in columns 1 to 6, while data in columns 7 to 15 have been derived from these measurements using the following factors: —

a) The density of air at N.T.P. is 1.292 gram per litre.
b) Air contains 0.031 % by volume of carbon dioxide and the natural radio-carbon background is 7.1×10^6 atoms per gram of air [Hagemann et al. 1959).
c) The half-lives of tritium and radiocarbon are 12.23 yrs and 5,600 yrs respectively.

The majority of the results are consistent but some variations are apparent and can be explained on known experimental observations.

On the 1956 and 1957 flights the liquid nitrogen caused cold spots on the sampler surface, which led to condensation of tropospheric water on the equipment. Humidity mixing ratios and tritium/hydrogen ratios were thus invalidated. This did not effect the actual tritium count since the tritium content of this tropospheric water was so low that its contribution could be neglected. If an average value of 0.041 g/Kg is taken for the humidity mixing

Table

Tritium, radiocarbon and humidity measurements

1	2	3	4	5	6	7	8
Date of Flight	Altitude (10^3 feet)	Collected CO_2 (cc. NTP)	Collected Water (mgm)	Tritium Activity (dpm)	C^{14} Activity (dpm)	Tritium atoms (10^9)	C^{14} atoms (10^{10})
20.9.56	93	33.2	800	252	—	2.33	—
10.5.57	92	20.4	180	53	—	0.49	—
2.5.58	89	118.3	21.3	239	2.4 ±1.0	2.21	1.0 ±0.4
13.5.59	91	95.0	17.8	223	9.9 ±0.8	2.06	4.2 ±0.3
5.6.59	80	150.1	5.6	338	3.0 ±1.7	3.13	1.3 ±0.7
15.6.59	93	69.0	14.1	283	7.0 ±1.8	2.62	3.0 ±0.8
18.6.59	90	157.9	50.6	397	2.8 ±0.4	3.68	1.2 ±0.2
7.7.60	89	139.6	19.6	451	2.2 ±0.5	4.18	0.9 ±0.2
11.9.60	99	76.5	11.1	203	0.65±0.21	1.88	0.3 ±0.1

ratio on these flights, the tritium/hydrogen ratios are 6×10^{-12} and 2×10^{-12} and thus fall within the range of the later results.

The highest humidity mixing ratio and correspondingly the lowest tritium/hydrogen ratio, were obtained on the flight of 18th June, 1959. Here the balloon was observed to rise through thin stratocumulus cloud and it is possible to attribute these results to the collection of tropospheric water on the surfaces of the sampling apparatus.

The lowest humidity mixing ratio and highest tritium/hydrogen ratio are both associated with the flight of 5th June, 1959, which was made at an altitude of 80,000 ft., i.e. at least 9,000 ft. lower than the other flights. This can be construed as evidence in favour of pick-up of tropospheric water on all the other flights, but this is considered unlikely since the flight was identical to the others except that the balloon was somewhat smaller.

Combining the sources of error associated with sample condensation, volume and counting measurements and calculations of the amount of air sampled it is estimated that the tritium concentrations, humidity mixing ratios and tritium/hydrogen ratios should be accurate to within 10 %.

In the case of radiocarbon concentrations, the errors in the activity determinations are the predominant factors. Since these factors vary from sample to sample the errors are quoted on the results in Table 2.

3.2 *Mass Spectrometric Determinations*

Results for the atom per cent carbon-13, oxygen-18 and deuterium in the carbon dioxide and hydrogen respectively are given in Table 3. Within the accuracy of the analyses, the atom percent deuterium was the same for high altitude moisture as for that at ground level. No increase was noted in the 1960 flights when heavy water (D_2O) was added to the balloon hydrogen. It was estimated that the balloon hydrogen contained approximately 6 g. of water and this amount of heavy water was added to give 50 atom percent deuterium, in the combined water. Since the deuterium figures on the water condensed in the sampling equipment showed no increase in these cases, moisture pick up from the balloon could be neglected.

The deuterium figures were in all cases higher than the value of 0.015 quoted for natural samples. The reason was that at the pressures pertaining in the mass spectrometer some $(H\,H\,H)^+$ ions were formed (Barnerd, 1953) and appeared at the HD peak so enhancing it. When a series of measurements were taken at various pressures, the plot of atom percent deuterium against pressure gave a value of 0.015 when extrapolated to zero pressure.

Carbon-13 and oxygen-18 results were the same for carbon dioxide from low level as for the high level samples and consistent figures were obtained over the whole series.

2

on samples of stratospheric air.

9	10	11	12	13	14	15
Weight of air sampled (g)	Mixing Ratio (g. water per Kg air)	Tritium Atoms (10^6 per g. of air)	C^{14} atoms (10^7 per g. of air)	C^{14} atoms (10^7 per g. of air) corrected for natural background	Tritium to Hydrogen atom ratio (10^{-12})	Atoms of excess C^{14} per atom of tritium
138	—	16.9	—	—	—	—
85	—	5.78	—	—	—	—
493	0.044	4.48	2.1±0.9	1.4±0.9	1.52	3.1±2.2
396	0.045	5.20	10.6±0.9	9.9±0.9	1.73	19.0±2.5
626	0.009	5.00	2.0±1.2	1.3±1.2	8.39	2.6±2.4
288	0.049	9.10	10.3±2.7	9.6±2.7	2.77	10.5±3.0
658	0.077	5.59	1.8±0.3	1.1±0.3	1.12	2.0±0.6
582	0.034	7.18	1.6±0.3	0.9±0.3	3.26	1.3±0.5
319	0.035	5.89	0.9±0.3	0.2±0.3	2.60	0.3±0.5

4. Discussion

Fig. 2(a) shows the time variation of the tritium concentration over the sampling period. Also plotted are the only other published tritium measurements above 80,000 feet at comparable latitudes to England (HAGEMANN ET AL., 1959); these were obtained over Minneapolis, U.S.A., using a somewhat different technique. The figure near each point refers to the height of the sample in thousands of feet. It is seen the agreement between the two sets of measurements is reasonable. The fluctuations in the tritium concentration up to 1958 during a time of many thermonuclear explosions, are not surprising. However, the 1959 and 1960 results obtained after the termination of weapon testing in November 1958, demonstrate no decline in the tritium concentration.

It is not possible to give an accurate estimate of what proportion of this tritium is due to natural cosmic ray sources. However, an order of magnitude estimate, using the natural production rate of about 1 tritium atom /cm^2/sec (WILSON & FERGUSSON, 1960), can be made if a very idealised stratosphere is considered. Assuming that two-thirds of the natural production takes place in a well mixed stratosphere

Table 3.

Mass spectrometric determinations of the atom % C^{13}, O^{18} and D on the stratospheric CO_2 and hydrogen and comparable results on samples collected at ground level.

Date of Flight	Carbon Dioxide		Hydrogen	Amount of D_2 added to balloon hydrogen (g)
	Atom % C^{13}	Atom % O^{18}	Atom % D	
13.5.59	1.165±0.001	0.208±0.001	0.0220±0.0005	Nil
5.6.59	1.159±0.001	0.209±0.001	0.0220±0.0010	Nil
15.6.59	1.167±0.001	0.209±0.001	0.0187±0.0007	Nil
18.6.59	1.167±0.007	0.209±0.002	0.0207±0.0006	Nil
7.7.60	1.163±0.004	0.210±0.001	0.0184±0.0010	6.0
11.9.60	1.168±0.004	0.210±0.001	0.0174±0.0010	6.5
Samples collected at ground level at the launching site 1959	1.165±0.007	0.208±0.001	0.0220±0.0013	
1960	—	—	0.0194±0.0010	

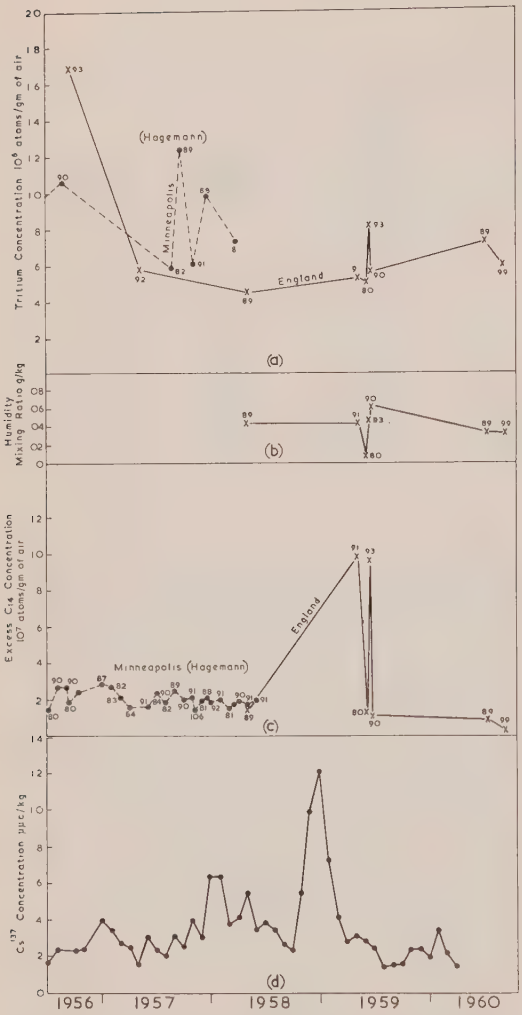


Fig. 2. (a) The tritium concentration, (b) the humidity mixing ratio, and (c) the excess carbon-14 concentration above 80 000 feet measured during the years 1956 to 1960. (d) The monthly mean concentrations of caesium-137, produced in nuclear tests, at a height of 47,000 feet over England during the same period.

in which the tritium has a residence time of 5 years, then the natural tritium concentration will be about 3×10^5 atoms/gm. This compares with the measured values of between 4×10^6 and 2×10^7 atoms/gm reported in this paper. It is probably fair to conclude that most of the tritium at present in the stratosphere is not of natural origin.

The humidity and excess carbon-14 variations are similarly presented in figs. 2(b) and 2(c). The more extensive carbon-14 results of HAGEMANN ET AL., (1959), are also plotted. The two peak carbon-14 concentrations found in 1959 are much greater than any variations reported in the American work, although the other results are in satisfactory agreement. Only the peak at 93,000 ft. on 15th June, 1959, is associated with a tritium peak. The correlation found between tritium and carbon-14 is not as good as that reported by Hagemann et al., but this may be due primarily to our carbon-14 counting being somewhat poor statistically, because of the small amounts of carbon dioxide collected.

The extensive measurements (BANNON ET AL., 1952, MURGATROYD ET AL., 1955, HELLIWELL ET AL., 1957), using aircraft fitted with the "Dobson-Brewer" frost-point hygrometer (BREWER ET AL., 1948), of the humidity up to 50,000 feet show that it is invariably dry there, with a humidity mixing ratio of about 0.002 g/Kg with no significant seasonal variations. The humidity data, presented in this paper, means that the air at heights of about 90,000 feet is very much moister with a mixing ratio of about 0.04 g/Kg.

The one result at 80,000 feet on 5.6.59 is the driest (.009 g/Kg) and may mean that the transition from moist to dry air on this occasion was in the region of this height.

The 1958 humidity result has already been compared with other data on the stratospheric humidity (BARCLAY ET AL., 1960). At that time the main evidence in support of a moist middle stratosphere was the occurrence of Mother-of-Pearl clouds at heights of about 80,000 feet over northern latitudes (HESSTVEDT, 1959). Automatic hygrometer balloon flights of BARRETT ET AL., (1950) also were in reasonable agreement, but this work also showed the lower stratosphere to be moist.

However, since the publication of the 1958 result, not only have the subsequent flights, presented here, confirmed the early result, but considerably more evidence in support of the presence of moist air above the dry lower stratosphere has been accumulating.

Recordings of the sun's infra-red spectrum by a spectrometer carried to heights of 92,000 feet by balloon (MURCRAY ET AL., 1960) over New Mexico, U.S.A., have been interpreted to point to this type of humidity profile in

the stratosphere. Similar work, using a spectrometer, mounted on a U.2 aircraft flown at heights of 65,000 feet, have yielded similar information on the vertical distribution of water in the stratosphere (Lab. of Astrophys., 1960). Recent soundings up to 100,000 feet, using a balloon borne automatic frost-point hydrometer, are again consistent with the concept of a dry lower stratosphere with moister air above (Mastenbrook & Dinger, 1960).

Brewer (1949) explained the low humidities measured in the lower stratosphere over England by postulating a meridional circulation such that air rises through the equatorial tropopause, moves horizontally polewards and finally sinks in the middle and high latitudes. The cold regions near the equatorial tropopause act as a cold trap which dries this circulating air. Dobson (1956) accounted for seasonal variations in the total ozone over temperate and polar regions, by suggesting that the cold pool over the winter pole stores ozone-enriched air which arrives by the meridional circulation from ozone producing regions. This air subsides into the lower stratosphere of the polar and temperate latitudes in the late winter and spring. This subsidence of the ozone rich air accounts for the spring peak in total ozone at these latitudes. The spring maximum of fission product fallout was first attributed to the same circulation by Stewart et al. (1957). The peaks in the radioactive debris of stratospheric origin falling out at temperate latitudes (Stewart et al., 1957, Crooks et al., 1960) were explained by a combination of this meridional circulation and air transfer from stratosphere to troposphere of those latitudes.

The measurements presented in this paper are not in accordance with this "Dobson-Brewer" stratospheric model. The tritium concentration during the period 1959 to 1960 indicates that the air at 90,000 feet is not exchanging with the troposphere at any appreciable rate. This would be consistent with the "Dobson-Brewer" circulation only if that circulation went to great heights, and took a period of several years. However, the most recent estimates of a stratospheric half residence time for radioactive debris of 1—2 years, Feely (1960), Peirson et al., (1960), are not in accordance with this latter requirement. The high humidities found at 90,000 feet are also quite inconsistent with the former requirement.

The source of the moist air above about 80,000 feet cannot be the tropical tropopause regions because of the low temperature there. The late winter and spring polar zones are also doubtful sources, since the ozone data is consistent with sinking air there. The most likely natural source, therefore, seems to be the summer polar upper troposphere which is comparatively warm. Humidity measurements above 40,000 feet at latitudes higher than 70° would be extremely interesting in this context.

Direct transfer of water into the stratosphere at 90,000 feet by thermonuclear explosions is not considered to be a likely cause of the high humidities. This would involve the injection of the order of 4×10^9 tons of water into the middle stratosphere, assuming the high humidities to be global in extent. An energy expenditure of about 2,000 Megatons TNT equivalent would be required to evaporate this quantity of water. This is an order of magnitude greater than the energy release of all weapon tests (U. K. *Min. of Defence*, 1959), only about a third of which was associated with water bursts.

In figure 2(d) the variations of the particulate caesium-137 concentration of the air at 47,000 feet over England, measured using aircraft fitted with filters, is reproduced from Peirson et al. (1960). There is little correlation between the caesium-137 variations at 47,000 feet and those of the tritium at 90,000 feet over the same period. It may be that the factors governing the source and transfer of air in the lower stratosphere over England are quite different from those pertaining at 90,000 feet. The humidity data are certainly very strong evidence in favour of this thesis.

5. Conclusions

The humidity and tritium measurements presented in this paper indicate that the large scale meridional circulation, suggested by Brewer, does not reach heights of 90,000 feet over England.

The world-wide distribution in the stratosphere of late 1959 of the particulate tracer tungsten-185, which was injected near the

equator in mid 1958, is also difficult to reconcile with the requirements of the Brewer circulation. FEELY and SPAR (1960) suggest that lateral mixing, with the possibility of some gravitational settling, within the stratosphere best explains the tungsten-185 distribution. However, it is difficult to use this explanation to account for the presence and maintenance of high humidities at 90,000 feet compared with 50,000 feet.

It would seem, therefore, that neither of these two mechanisms of meridional transfer within the stratosphere can fully account for all the information available on the distribution of the various tracers (water, ozone, radioactive debris) found at those heights. In a subsequent paper, two of us (P. G. and F. B.) will consider this aspect in more detail.

6. Acknowledgements

The authors thank W. P. Cook, J. Davies and H. Heitler for controlling the balloon launches, D. E. Watt for advice on the carbon-14 counting, D. E. Lloyd for work on the radio control system, and Miss N. Campbell who carried out the Mass Spectrometric analyses. Drs. F. Morgan and J. V. Jelley are also thanked for their interest and encouragement.

The Radar tracking facilities were provided by O.C. R.A.F. Fighter Command and W. G. Harper of the Meteorological Office.

REFERENCES

BANNON J. K., FRITH R., and SHELLARD H. C., 1952: Humidity of the upper troposphere and lower stratosphere over southern England, *Meteorological Office, Geophys. Memoirs* No. **88**.

BARCLAY F. R., ELLIOTT M. J. W., GOLDSMITH P., and JELLEY J. V., 1960: A direct measurement of the humidity in the stratosphere using a cooled vapour trap, *Quart. J. R. Met. Soc.*, **86**, p. 259.

BARNERD G. P., 1953: *Modern Mass Spectrometry*, Inst. of Phys. London.

BARRETT E. W., HERNDON L. R., and CARTER H. J., 1950: Some measurements of the distribution of water vapour in the stratosphere, *Tellus*, **2**, p. 302.

BREWER A. W., 1949: Evidence for a world circulation provided by the measurements of helium and water vapour distribution in the stratosphere, *Quart. J. R. Met. Soc.*, **75**, p. 351.

BREWER A. W., CWILONG B. M., and DOBSON C. M. B., 1948: Measurement of absolute humidity in extremely dry air, *Proc. Phys. Soc.*, London A. **60**, p. 52.

CARSLAW H. S. and JAEGER J. C., 1947: *Conduction of Heat in Solids*, Oxford, p. 175.

CROOKS R. N., OSMOND R. G. D., FISHER MISS E. M. R., OWERS M. J., and EVETT T. W., 1960: The deposition of fission products from distant nuclear test explosions. Results to the middle of 1960, *A.E.R.E. Report 3349*.

DOBSON C. M. B., 1956: Origin and distribution of the polyatomic molecule in the atmosphere, *Proc. Roy. Soc. A.*, **236**, p. 187.

FEELY H. W., 1960: Strontium-90 content of the stratosphere, *Science*, **131**, p. 645.

FEELY H. W., and SPAR J., 1960: Tungsten-185 from nuclear bomb tests as a tracer for stratospheric meteorology, *Nature*, **188**, p. 1062.

GARD G. A., 1958: An estimation of tritium activity in a hydrogen sample, *A.E.R.E. N/M 86*.

GLUECKAUF E., 1944: Carbon dioxide content of atmospheric air, *Nature*, **153**, p. 620.

GOLDSMITH P., JELLEY J. V., BARCLAY F. R., ELLIOTT M. J. W., and OSBORN A. R., 1960: Some preliminary measurements of the tritium and carbon-14 content of the stratosphere over England, *A.E.R.E. Report 3271*.

GOLDSMITH P. and PARHAM A G., 1960: An apparatus for sampling the stratosphere, s *A.E.R.E. Report 3272*.

HAGEMANN F., GRAY J., MACHTA L., and TURKEVICH A., 1959: Stratospheric carbon- 14 carbon dioxide, and tritium, *Science*, **130**, p. 542.

HELLIWELL N. C., MACKENZIE J. K., KERLEY M. J., 1957: Some further observations from aircraft of frost point and temperature up to 50,000 ft., *Quart. J. R. Met. Soc.*, **83**, p. 257.

HESSTVEDT E., 1959: Mother of pearl clouds in Norway, *Geofysiske Publikasjoner Geophys. Norvegica*. Vol. **XX**, No. 10.

LAB. OF ASTROPHYS. AND PHYS. METEOR., J. HOPKINS UN., 1960: *Prelim. report on solar observations from an U-2 observatory*.

MASTENBROOK H. J. and DINGER J. E., 1960: The measurement of water-vapour distribution in the stratosphere, *N.R.L. Report 5551*.

MURCRAY D. G., MURCRAY E. H., WILLIAMS W. J., and LESLIE F. E., 1960: Water distribution above 90,000 feet, *J. of Geophys. Res.* **65**, p. 3641.

MURCRAY D. G., MURCRAY E. H., and WILLIAMS W. J., 1960: Variation of infrared solar spectrum with altitude, *Un. of Denver Final Report Contact AF 19(604)-2049*.

MURGATROYD R. J., GOLDSMITH P., and HOLLINGS W. B. H., 1955: Some measurements of humidity from aircraft up to heights of about 50,000 feet over southern England, *Quart. J. R. Met. Soc.*, **83**, p. 533.

PEIRSON D. H., CROOKS R. N., and FISHER E. M. R. 1960: Radioactive fall-out in air and rain, *A.E.R.E. Report 3358*.

STEWART N. G., OSMOND R. G. D., CROOKS R. N., FISHER MISS E. M. R., 1957: The world-wide deposition of long-lived fission products from nuclear test explosions, *A.E.R.E. Report 2354*.

UNITED KINGDOM MINISTRY OF DEFENCE, PRESS RELEASE 5TH MAY, 1959: *Information on fission yield from the United Kingdom, U.S. and Soviet nuclear test events*.

WILSON A. T. and FERGUSSON G. J., 1960: Origin of terrestrial tritium, *Geochim. et Cosmochim. Acta*. **18**, p. 273.

Forecasting Thunderstorms and Showers by the Slice Method

By N. S. SHISHKIN, Main Geophysical Observatory, Leningrad, U.S.S.R.

(Manuscript received December 16, 1960)

Abstract

The author gives a further development of the slice method proposed by J. Bjerknes in 1938. The formulae obtained use data from aerological soundings to compute the vertical thickness of convective clouds, their optimum amount, and the velocity of vertical development. The paper deals with the way of forecasting thunderstorms and showers based on the slice method.

Introduction

The slice method developed by J. Bjerknes (1938), S. Petterssen (1939) and A. F. Dubuk (1945), permits computation of the energy resources of instability in the atmosphere spent on convection. Unlike the particle method, the slice method applied to the development of convective clouds takes into account both the amount of energy associated with the ascent of cloud air and the amount of energy spent on the compensating descent of air outside the clouds.

N. Beers (1945) proposed a method of forecasting thunderstorms and showers based on the slice method. The method was tested by S. Petterssen and others (1946) in England and by S. Banerji (1950) in India.

Computations were based on the determination of mean kinetic energy per unit air mass participating in convection. Computations were made separately for dry and saturated slices, the saturated slices being defined as those with relative humidity of 70 % or more. However, the methods of computation were too inaccurate and the method of forecasting proposed by N. Beers gave unsatisfactory results.

The author proposes a new method of forecasting thunderstorms and showers on the basis of a certain development of the slice method Shishkin (1957). The method has been tested by many forecasting units of the Soviet Union and showed a high percentage of accuracy for the air mass thunderstorms and showers.

Theoretical Foundations of the Slice Method

J. Bjerknes (1938) established thermodynamic conditions for the development of convective clouds. He discussed an atmospheric layer containing clouds and cloudless air. The change in the heat content of the layer per unit time during the development of clouds is expressed by the formula

$$\frac{\partial Q}{\partial t} = C_p \left(M_d \frac{\partial T_d}{\partial t} + M_s \frac{\partial T_s}{\partial t} \right) \qquad (1)$$

where C_p is the specific heat of air at constant pressure, M_d and M_s – masses of dry and saturated air participating in circulation, T_d and T_s – temperatures of dry and saturated air, respectively. The right hand side of expression (1) may be transformed if we add and subtract the term $M_s \frac{\partial T_d}{\partial t}$. Then we obtain

$$\frac{\partial Q}{\partial t} = C_p \left[(M_d + M_s) \frac{\partial T_d}{\partial t} + M_s \left(\frac{\partial T_s}{\partial t} - \frac{\partial T_d}{\partial t} \right) \right] \qquad (2)$$

The first term in square brackets gives the change of heat energy in the slice due to the general warming (or cooling), while the second term shows excessive heating (or cooling) of cloud air, i.e. the heat energy of circulation.

When we take into account the law of conservation of mass

$$M_d v_d + M_s v_s = 0 \qquad (3)$$

and make a number of simplifying assumptions, we obtain the following expression for the change of heat energy of convection when the cloud air ascends through a slice ΔH with lapse-rate γ (SHISHKIN, 1959):

$$\Delta Q = C_p \bar{\varrho}\, S_s \left[(\gamma - \bar{\gamma}_s) - \frac{S_s}{S_d} (\gamma_d - \gamma) \right] (\Delta H)^2 \qquad (4)$$

or, for a unit mass of cloud air,

$$\Delta q = C_p \left[(\gamma - \bar{\gamma}_s) - \frac{S_s}{S_d} (\gamma_d - \gamma) \right] \Delta H \qquad (5)$$

In the above formulae $\bar{\gamma}_s$ is the mean value of saturated lapse-rate in the slice ΔH, γ_d – dry-adiabatic lapse-rate, S_s and S_d – amounts of clouds and cloudless air, respectively (10 tenths are considered a unit), $\bar{\varrho}$ – mean density of air in the slice. It is assumed that the lapse-rate γ is within the range $\bar{\gamma}_s < \gamma < \gamma_d$.

The right hand sides in formulae (4) and (5) are positive if the following condition is fulfilled:

$$\frac{S_s}{S_d} \leqslant \frac{\gamma - \bar{\gamma}_s}{\gamma_d - \gamma} \qquad (6)$$

The inequality (6) is called the Bjerknes criterion. It gives the condition for the development of convective clouds.

Let us now compute the work done when the development of convective clouds occurs in a slice with a lapse-rate γ.

Fig. 1 shows schematically a small portion of the slice containing one convective cell. Here h is the vertical thickness of a cloud at a time t. When ascending, the cloud mass becomes warmer than the surrounding stable air. The ascending force related to the unit mass at the height z is:

$$f_s = g \frac{T_s - T}{T} \qquad (7)$$

where g is acceleration due to gravity, $T_s = T_0 - \gamma_s z$ is the temperature of the rising cloud air, $T = T_0 - \gamma z$ is the temperature of stable air at the height z, T_0 is the air temperature at the lower boundary of the slice. It is assumed that the air ascends adiabatically.

Substituting values T_s and T into (7) we obtain

$$f_s = \frac{g (\gamma - \gamma_s) z}{T_0 - \gamma z} \qquad (8)$$

By the time the upper boundary of the cloud reaches the height h, the following amount of cloud air will pass through the horizontal plane at the height z:

$$\Delta M_s = \varrho_s S_s (h - z)$$

where ϱ_s is the density of cloud air.

In accordance with the law of continuity, the same amount of dry air must descend. The thickness of the slice of dry air that has descended, Δz_d, may be determined from the condition

$$\varrho_s S_s (h - z) = \varrho_d S_d \Delta z_d$$

where ϱ_d is the density of dry air. Neglecting the difference between the densities of cloud air and dry air we obtain

$$\Delta z_d = \frac{S_s}{S_d} (h - z)$$

Let us now consider the way by which the compensating descent of dry air takes place. When the cloud air ascends due to buoyancy, the dry air situated above the cloud is forced to move upward. If the dry air ascends adiabatically it becomes colder than the stable air at the corresponding level. The motion is retarded and the ascending air is spreading out. Being colder, the dry air displaced aside begins to de-

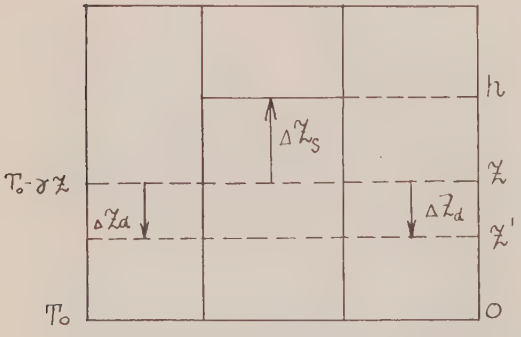


Fig. 1. Scheme of development of the convective cloud.

scend in the cloudless space. In its turn, the ascent of cloud air in the lower part of the slice promotes the development of a descending flux in the surrounding dry air. Thus a closed circulation appears.

If we do not take into account turbulent mixing of the air and other non-adiabatic processes, then the dry air which descends near the upper boundary of the cloud will return to its initial level, where it began ascending, with its same initial temperature. Thus, we may consider that the dry air which descends from a level z has an initial temperature $T = T_0 - \gamma z$.

At the level $z^1 = z - \Delta z_d$ when the descent is dry-adiabatic, the air has the temperature $T_d = T_0 - \gamma z + \gamma_d \Delta z_d$. The buoyancy force applied to a unit mass of dry air at the level z^1 (in relation to the stable air) will be therefore

$$f_d = g\,\frac{T_d - T}{T} = g\,\frac{(\gamma_d - \gamma)\,\Delta z_d}{T_o - \gamma\,(z - \Delta z_d)} \qquad (9)$$

This force, as well as the buoyancy force for the saturated air, is directed upward.

Comparison of formulae (8) and (9) shows that the correlation of buoyancy forces is different at different levels. Having made an equation of these expressions, we shall find the level where the buoyancy forces for the cloud and dry air are equal to one another:

$$Z_K = \frac{h}{1 + \dfrac{S_d}{S_s}\,\dfrac{\gamma - \gamma_s}{\gamma_d - \gamma_s}} \qquad (10)$$

We have $f_s > f_d$ above the critical level and $f_s < f_d$ below the critical level, if we assume that during the development of circulation the temperature at the lower boundary of the cloud remains constant. Thus, in the lower part of the slice retardation of circulation takes place during the development of the cloud.

However, it is not the relation between f_s and f_d that determines the development of circulation as a whole. The motion is determined by the integral action of buoyancy forces of all the masses participating in the circulation. The integral sums of the buoyancy forces affecting the cloud and dry air are:

$$F_s = \int f_s\,dM_s \qquad , \qquad F_d = \int f_d\,dM_d$$

Substitution of expressions (8) and (9) into the above expressions gives on integrating:

$$F_s = \frac{g}{2T_0}\,(\gamma - \bar{\gamma}_s)\,\bar{\varrho}\,S_s\,h^2\,,$$

$$F_d = \frac{g}{2T_0}\,(\gamma_d - \gamma)\,\bar{\varrho}\,S_s\,h^2 \qquad (11)$$

The work done by buoyancy forces in forming clouds ΔH thick is:

$$\Delta A = \int F_s\,dh_s + \int F_d\,dh_d$$

Since $dh_d = -\dfrac{S_s}{S_d}\,dh_s$ we obtain on integrating:

$$\Delta A = \frac{g\bar{\varrho}S_s}{6T_0}\left[(\gamma - \bar{\gamma}_s) - \frac{S_s}{S_d}\,(\gamma_d - \gamma)\right](\Delta H^3) \qquad (12)$$

Comparing this with (4) and taking account of $g \approx C_p\gamma_d$ we obtain the formula:

$$\Delta A = \frac{T_0 - T_d}{6T_0}\,\Delta Q \qquad (13)$$

as $\gamma_d \Delta H = T_0 - T_d$, where T_d is the temperature at the upper boundary of the layer of thickness ΔH, when the ascent is dry-adiabatic.

Consequently, the value

$$\eta = \frac{T_0 - T_d}{6T_0} \qquad (14)$$

is the coefficient of transformation of heat energy into kinetic energy. It has been obtained under the assumption that at the given level the vertical velocity of the cloud is uniform. The same assumption is used for the velocity of the compensating motion of dry air.

If we assume that the distribution of velocities of vertical motions in convective clouds and in the environment is approximately the same as with regulated convection, we may then use the following relation: mean density of heat flux in a convective cell is equal to 1/6 of density of heat flux in the centre of the cell (SHISHKIN, 1954).

Thus, if the ascending flux is the same along all its cross-section as in the central part of the cells and if the velocity of the descending flux be correspondingly greater, then the work done would be six times greater.

The maximum coefficient of transformation of heat energy of convection into kinetic energy is:

$$\eta = \frac{T_0 - T_d}{T_0} \qquad (15)$$

Let us now find the velocity of development of convective cloudiness.

The amount of work done in developing the convective cloud ΔH thick may be considered equal to the kinetic energy of circulation:

$$E_{KIN.} = M_s \frac{v_s^2}{2} + M_d \frac{v_d^2}{2}$$

Using the continuity equation (3) we obtain:

$$v_d = -\frac{M_s}{M_d} v_s$$

If we assume the clouds to be of cylindrical shape and make simple transformations we obtain the expression:

$$E_{KIN.} = \bar{\varrho} \frac{S_s}{S_d} \frac{v_s^2}{2} \Delta H \qquad (16)$$

Equating the right hand sides of (12) and (16) we obtain the values for the kinetic energy of unit cloud mass at the moment when the thickness of the cloud reaches the value ΔH:

$$\frac{v_s^2}{2} = \frac{gS_d}{6T_0}\left[(\gamma - \bar{\gamma}_s) - \frac{S_s}{S_d}(\gamma_d - \gamma)\right](\Delta H)^2 \qquad (17)$$

Equation (17) leads to the condition of development of convective clouds (6) obtained by J. Bjerknes.

Allowing for the energy expended in lifting dry air above developing clouds, the condition of appearance of convective clouds becomes

$$\gamma \geqslant \frac{\gamma_s + \gamma_d}{2} \qquad (18)$$

For practical purposes it is convenient to introduce a substitution in (17):

$$(\gamma - \bar{\gamma}_s)\Delta H = T_s - T, \quad (\gamma_d - \gamma)\Delta H = T - T_d,$$

where T is the temperature of stable air at the level ΔH. T_s and T_d are temperatures of the dry- and saturated-adiabatically ascending air, respectively, when the ascent occurs from the lower boundary up to the level ΔH.

Considering $S_d = 1 - S_s$ and replacing $g = C_p\gamma_d$, we ultimately obtain:

$$\frac{v_s^2}{2} = C_p\left[(T_s - T) - S_s(T_s - T_d)\right]\frac{T_0 - T_d}{6T_0} \qquad (19)$$

When the cloud is developing through several atmospheric layers with different values of the lapse-rate γ_k, it is necessary to sum up the changes of heat energy for the unit cloud mass:

$$\Delta q_K = C_p\left[(T_s - T)_K - S_s(T_s - T_d)_K\right]$$

and to multiply the result by the coefficient of transformation of heat energy into kinetic energy:

$$\eta_n = \frac{T_0 - Td_n}{6T_0}$$

where Td_n is the air temperature at the upper boundary of the layer n when dry-adiabatic ascent from the condensation level occurs.

We obtain:

$$\Delta\left(\frac{v_s^2}{2}\right) = \frac{T_0 - T_{dn}}{6T_0}\sum_{K=I}^{n} C_p\left[(T_s - T)_K - S_s \cdot (T_s - T_d)_K\right] \qquad (20)$$

This is an approximate formula since it does not take an accurate account of the correlation between temperatures of the ascending and descending air when convection develops through several atmospheric layers.

In order to compute the velocities of vertical motions in clouds it is necessary to make the sums by slices beginning from the condensation level. The conditions for development of clouds through several atmospheric layers are most favourable in the case when maximum amount of kinetic energy is guaranteed.

In order to find out the corresponding amount of cloud (which may be termed the optimum amount), we have to determine the variation of the work done, given by equation (12), as S_s varies.

Assuming the derivatives to be zero we obtain the optimum amount of clouds:

$$S_0 = 1 - \sqrt{\frac{\overline{T - T_d}}{\overline{T_s - T_d}}} \qquad (21)$$

When convective clouds develop through several atmospheric layers,

$$S_0 = 1 - \sqrt{\frac{\sum_K \overline{(T - T_d)_K}}{\sum_K \overline{(T_s - T_d)_K}}} \qquad (22)$$

Substitution of $S_s = S_0$ into (20) gives the velocities of the ascending air motions in convective clouds under the optimum condi-

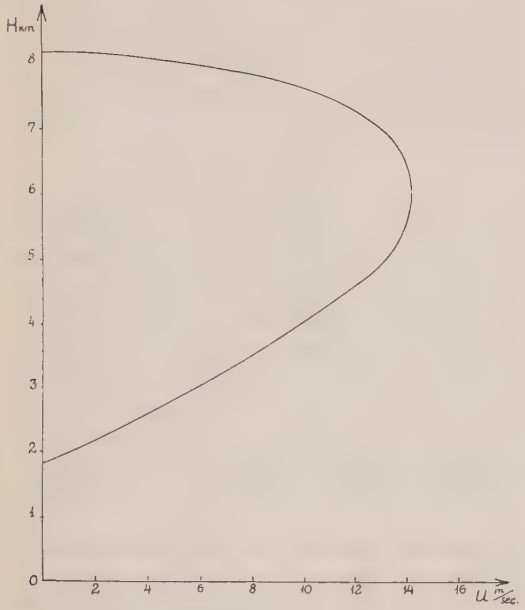


Fig. 2. Change of velocity of ascending flux in clouds with altitude on Juni 6, 1959, after computation data.

tions of development. Fig. 2 shows the computed change of velocities of the ascending motions in convective clouds with height. The data used refer to the morning radio-souding in Ruispiri (East Georgia, U.S.S.R.) on June 6th, 1959. We can see that the velocity could reach the value of 14 m/sec at the height of 6 km.

The optimum amount of clouds varies for clouds of different vertical thickness, i.e. for clouds in different stages of development.

In the case when it changes little with height, conditions are favourable for the continuous development of the same convective clouds. When S_0 changes with height by a factor of two or more, the development of some clouds promotes the dissipation of others. Therefore, under such conditions, cloud growth proceeds in a rather unstable manner: some clouds are intensively developing while others are rapidly dissipating.

With an unstable atmospheric stratification, the velocity of descending motions in those convective clouds which are dissipating, may also be computed using formula (20). In this the initial value of T_0 should be taken as the temperature at the upper boundary of the dissipating clouds. The computation of velocity is carried out by slices from the cloud top to the cloud base.

Computation of Vertical Thickness of Convective Clouds

Two necessary conditions for the development of convective clouds are the accumulation of energy of atmospheric instability and sufficient humidity. The energy released by convection may be computed by summing the expression on the right hand side of formula (20) for each slice. Different empirical rules are used to evaluate the moisture supply; some of them will be treated later on.

The amount of convective clouds is a value that changes greatly with time. With sufficient air humidity it depends mainly on the lapse-rate in a layer, several hundred meters thick, above condensation level. In order to solve the problem of forecasting convective cloudiness in the most general way, we may compute the cloud thickness by setting the amount of cloud in (20) equal to 1, 2, 3 . . . tenths, i.e. by giving S_s the values 0.1, 0.2, 0.3 . . . Thus we can find the possible thickness of clouds for different cloud amounts, given the temperature stratification.

Let us now consider the problem of forecasting the diurnal variation of convective cloud from the morning aerological sounding.

The basic data for the computation are as follows: (1) data of aerological sounding[1], (2) daily forecast of maximum temperature at the ground. The forecast of dew point for the period of maximum development of clouds is also desirable. When it is not available, forecasters may use the actual value of the dew point near the ground at the moment of forecasting or the maximum value of the dew point (taken, as a rule, either at the ground or at the upper point of ground inversion) from the data of the morning aerological sounding.

The condensation level is given by the point of intersection of the specific humidity curve corresponding to the chosen dew point and the dry-adiabatic curve passing through the forecast ground temperature.

[1] It should be considered that the aerological sounding made during rain or soon after it does not give representative data on the lapse-rate. It should not be used in this work.

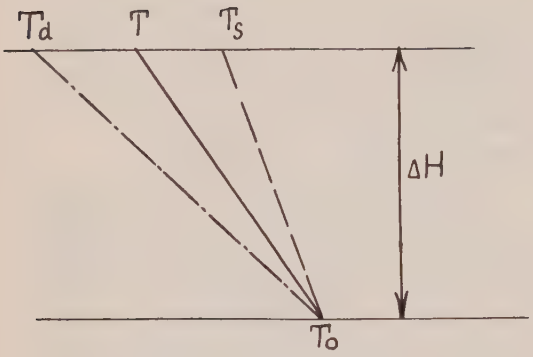


Fig. 3. Scheme of plotting on the aerological diagram.

If the point of intersection of these two curves lies to the left of the stratification curve by an amount exceeding some critical value (for the temperate climatic zone of the U.S.S.R. this critical value ranges from 3° C to 5° C), then the humidity is not sufficient to produce convective clouds. When humidity is sufficient, suitable slices for computation are chosen in accordance with points of inflection of the strafitication curve. Slices 50 – 100 mb thick prove to be the most expedient. To simplify the computation it is convenient to have an additional plotting on the aerological diagram: let us draw the dry and saturated adiabatic curves through the point on the temperature stratification curve at the slice base up to the slice top (see fig. 3). The values $(T_s - T)$ and $(T_s - T_d)$ can be measured off the diagram.

In order to determine the possible vertical thickness of clouds when their amount is S_s, we must sum up the expressions (20) from the condensation level to the height where the sum becomes zero[1]. This will be the level of the cloud upper boundary at maximum development.

Summing up the values $(T_s - T)$ by slices beginning with the condensation level, we obtain the greatest possible thickness of convective clouds when their amount is small $(S_s \to 0)$.

A comparision of computed maximum thicknesses of thunderstorm and shower clouds with a large number of the actually observed cases shows that the average computation error does not exceed 0.2—0.3 km.

For convective clouds of small thickness the difference between the computed and observed data is usually much smaller.

Forecasting of Thunderstorm and Showers

The computation data for vertical thickness of convective clouds may be used for the forecast of thunderstorms and showers.

According to the data of aircraft studies of convective clouds carried out over a few years at the Main Geophysical Observatory, U.S.S.R. the minimum vertical thickness of shower clouds for temperate zone summers is 2.2 km, while the minimum thickness of thunderstorm clouds is 4.0 km. We can determine empirically some critical values of the computed vertical thickness above which thunderstorms and showers may occur. The values may be different in various geographical conditions. Thus, for the zones of moderate and subtropical climate the critical values of thickness for shower clouds are likely to range between 2.2 km and 3.2 km[2] while for thunderstorm clouds they range between 4.0 km and 4.5 km. Even brief experience allows one to specify these criteria for particular districts.

Application of this method in the U.S.S.R. shows that the forecast of thunderstorms and showers is best justified for territories of 100 – 200 km radius. The forecast can be made for the day following the morning aerological sounding (or for a night after the evening sounding) performed at a station within the forecast territory. The range of forecasting may be extended to 1/2—1 day it we use the data of the evening or morning aerological soundings performed on the previous day at an upwind station within the same air mass.

Apart from the alternative forecast of thunderstorms and showers, the computation gives the portion of the territory where thunderstorms and showers may be expected. As a result of computation we know the amount of cloud when their thickness exceeds the critical values for thunderstroms and showers. It should be stressed that the results obtained do

[1] If the sum of terms in a slice is negative and in higher slices becomes positive again, then, as a rule, the upper boundary of clouds is considered the highest level where the sum becomes zero.

[2] If the computed upper boundary of convective cloudiness reaches —10°, then showers may be forecast for thickness of clouds less than 2.2 km.

not correspond to the amount of cloud seen from one station but give the relative proportion of the forecast area covered by clouds of super-critical depth. Thus, one tenth of cloudiness of thickness more than 4.5 km means that the thunderstorm cloudiness can develop over 10 % of the territory. In this case thunderstorms could be detected at 10 % of stations if there were no displacement of clouds. Forecasting experience shows that on account of cloud motion the calculated percentage of territory with thunderstorms and showers must be multiplied by an empirical coefficient 1.5—2.5 that varies in different geographical conditions. For the plains on the U.S.S.R. territory the coefficient is 2.0—2.5 while for the mountainous regions it is 1.5—2.0.

The mean error in forecasting the percentage of territory with thunderstorms and showers was $\pm$ 12 %.

The following empirical rules are used to take into account air humidity in the free atmosphere in the process of forecasting thunderstorms and showers:

1. Thunderstorms are not forecast if the total deficit of dew point[1] at the levels of 850, 700 and 500 mb exceeds some critical value (in temperate zone of the U.S.S.R. it various from 26° to 30°).

2. Showers are not forecast if the total deficit of dew point at the levels of 850 and 700 mb exceeds some critical value (in the above zone it varies from 18° to 20°).

If the test shows that the air humidity is small, computation may be omitted.

Test of the described method was made for a considerable part of the territory of the U.S.S.R. in 1956—57. During that period over 1,500 experimental forecasts were made. On the average, forecasts proved to be correct in 92 % of cases for air mass thunderstorms and showers, when the daily forecasts were based on data of the morning aerological soundings of the same day. For the extended forecasts the verification is somewhat poorer due to additional errors in forecasting air mass transfer and to neglecting air mass transformation. For the frontal thunderstorms and showers the verification of forecasts was 70—80 %.

Greatest success has been achieved for territories of the interior continental plains (up to 97 %). For the shores of seas and oceans the method gives good results when the air mass is transported from the land. When the air mass is transported from the sea, a considerable transformation of air during a day occurs and the method gives bad results. In mountainous regions the development of convective clouds has a number of peculiarities.

Not infrequently they develop over mountains even when the temperature stratification is stable in the lower layers of the air above an adjoining plain or a valley where clouds either do not develop at all or are of small thickness.

When forecasting, account of this phenomenon may be taken in the following way: if in a slice situated above the condensation level the lapse-rate is less than the saturated-adiabatic value ($\gamma < \gamma_s$), while higher the stratification becomes saturated-unstable, computation is made beginning with the lower boundary of the unstable layer. In this case the probable vertical thickness of convective clouds is found as the difference between the calculated upper boundary of clouds and the lower boundary of the unstable layer.

In mountainous regions the verification of forecasts is somewhat lower than it is for the plains (82—90 % for the air mass conditions).

Zones of thunderstorms and showers may be specified more accurately by interpolating values of maximum thickness after the data of several aerological stations.

Computation of one aerological sounding takes only 10—15 minuts. Thus, the proposed method may be applied in operational practice when forecasting thunderstorms and showers for the warmer part of the year.

[1] Deficit of dew point is the difference between actual air temperature and dew point.

REFERENCES

BANERJI, S., 1950: Forecasting of thunderstorms at Nagpur by slice method. *Ind. J. Meteorol. a. Geophys.* **1**, No. 3.

BEERS, N., 1945: *Atmospheric stability and instability.* Handbook of Meteorology, N.Y.

BJERKNES, J., 1938: Saturated-adiabatic ascent of air through dry-adiabatically descending environment. *Q. J. Roy. Met. Soc.* **64**, No. 275.

DUBUK, A. F., 1945: On the problem of realization of saturated instability and convection. *Trudy NIU GUGMS SSSR*, Ser. **1**, No. 7.

PETTERSSEN, S., 1939: Contribution to the theory of convection. *Geoph. publ.* **12**, No. 9.

— et al. 1946: Convection in theory and practice. *Geoph. publ.* **16**, No. 10.

SHISHKIN, N. S., 1954: Clouds, precipitation and thunderstorm electricity. GITTL, M.

— 1957: On forecasting of thunderstorms and showers by slice method. *Meteorol. and Hydrol.* No. 8.

— 1959: Studies of dissipation of convective clouds with the unstable stratification of the atmosphere. *Trudy GGO*, No. **102.**

Radiation to Actinometric Receivers in Its Dependence on Aperture Conditions

By ANDERS ÅNGSTRÖM, The Eppley Foundation for Research, Newport R.I., U.S.A.

(Manuscript received May 5, 1961)

Abstract

After an introduction on basic conceptions, concerning the influence of aperture conditions on the readings of actinometers, a short review is given on recent results as regards the circumsolar sky radiation. The combined effect of aperture and sky radiation on pyrheliometric readings, in the case of point receivers, is discussed. Conclusions are drawn as regards the influence of aperture conditions on (a) solar constant determinations and (b) on the difference between readings with aid of different pyrheliometers. The principles which ought to be applied in the construction of pyrheliometers for different kind of use, are briefly discussed.

1. When a pyrheliometer is pointed at the sun for measurement of the sun radiation, it takes in not only the direct radiation from the disk of the sun but also that from a portion of the sky in its immediate vicinity. The amount of this circumsolar sky radiation is dependent on the scattering properties of the atmosphere, and also on the geometrical construction of the receiver and of the pyrheliometer tube. The construction problem is often referred to as a question of the *opening angle* or the *aperture* of the instrument. With the aperture we understand, for an instrument with circular diaphragms, the top angle of the cone within which radiation from without is able to reach the central point of the receiver. The aperture can be defined through the radius R of the most narrow diaphragm, active in the instrument, and the distance (l) from its center to the center of the receiver (see fig. 1). If the aperture angle is denoted by φ, we have:

$$\operatorname{tg} \frac{\varphi}{2} = \frac{R}{l} \qquad (1)$$

If the instrument is directed towards a surface of uniform radiation intensity I, covering the whole aperture angle, the radiation reaching the central point of the receiver and referred to a unit of surface is given by

$$\pi I \sin^2 \frac{\varphi}{2} \qquad (2)$$

If we, with Dubois, define the "surface angle ratio" (the Flächenwinkelverhältnis) ψ, as the ratio between what a receiver, whose aperture is limited by diaphragms, measures and what is measured at an unlimited exposure to the whole hemisphere of the same instrument, the named ratio Ψ, in the case that the receiver is concentrated in a point, situated on the axis through the center of the diaphragm is, given by:

$$\psi = \sin^2 \frac{\varphi}{2} \qquad (3)$$

This is an ideal case which only seldom is realized in the actual actinometers used for field measurements. In general the receiver has

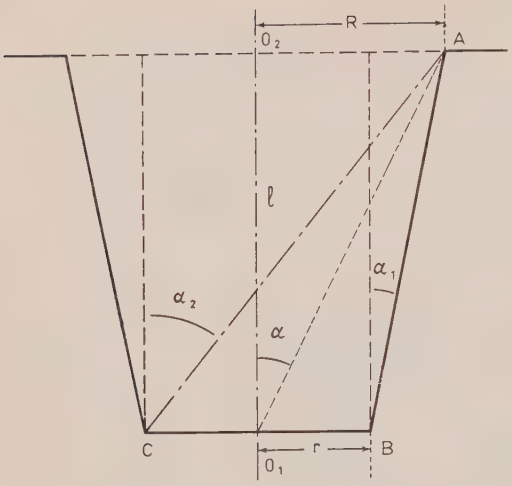


Fig. 1. Actinometer with circular diaphragm (schematic)
R = radius of diaphragm, r = radius of receiver,
l = distance between both α = opening angle $\left(\frac{\varphi}{2}\right)$
α_1 = slope angle α_2 = limit angle.

dimensions so large that the laws derived above are not strictly applicable. An important case, first treated by J. H. LAMBERT (1760) is the following.

We consider an actinometer with circular receiver and circular diaphragms, schematically represented in Fig. 1; R is here the radius of the diaphragm, r the radius of the receiver and l the distance between diaphragm and receiver. If the instrument is pointed at a source with uniform luminosity, I, the problem of finding the radiation falling on the receiver is that of determining the radiation exchange between two circular surfaces, parallel with one another and with their centers on an axis perpendicular to them both. Lambert derived for the radiation Q falling through the diaphragm on the receiver the following expression:

$$Q = \frac{\pi^2}{4} I (AC - AB)^2 \qquad (4)$$

where AC and AB have the meaning apparent from the figure. Using the geometrical relationships evident from fig. 1, we obtain from equation (4) for the average radiation Q_1, falling on the unit of surface of the receiver:

$$Q_1 = \pi I \frac{R^2}{\left(\frac{AC + AB}{2}\right)^2} = \pi I \psi \qquad (5)$$

Here evidently:

$$\psi = \frac{R^2}{\left(\frac{AC + AB}{2}\right)^2}$$

where ψ is the surface angle ratio. For an instrument with the receiver concentrated at a point, we have $AC = AB$, and the equation (5) then coincides with equation (2), which thus represents the limit conditions of (5).

The equations presented above have a *basic interest* when we consider the results of measurements with actinometers used for observations of the radiation from limited parts of a sky of comparatively uniform radiation power. The practical limitation of the applicability of the formula (4) of Lambert and of equation (5) derived from it, is due chiefley to the fact that in most practical cases the receiver has not a uniform sensitivity for the radiation falling upon it. A simple case is that of the silver-disk pyrheliometer where a flat cylindrical disk is heated by radiation, the temperature being measured at the central part of the disk. It is evident that, if we apply the usual prescribed procedure of reading the pyrheliometer and let a given amount of radiation fall on different parts of the disk, we will in general, if we apply a given constant to the readings, get different values for the radiation. The values will vary with the distance of the illuminated spot from the center. If we introduce a function $F(r)$ for the sensitivity in its dependence on the distance r from the center of the receiver, we are able to modify the law of Lambert, and derive formulas valid for actinometers of various construction. L. BOSSY and R. PASTIELS (1948, 1959) have, in papers of rather fundamental importance, treated this subject in a strict mathematical way and shown how the theory may be applied to practical cases. In many cases actually occurring, however, an approximation on the basis of the original formula of Lambert seems to give sufficient accuracy. LINKE (1931) has suggested the use of two parameters a and b, which he calls the *constants of normalization*, for defining the aperture conditions of actinometers with circular diaphragms. With the indications used above we have: $a = \frac{R}{r}$ and $b = \frac{l}{r}$, and consequently:

$$\frac{a}{b} = \frac{R}{l} = \operatorname{tg} \frac{\varphi}{2} \qquad (6)$$

Linke's constants of normalization may, if R and r, and R_1 and r_1 are used to denote the half length and half width of rectangular diaphragms, respectively receivers, be applied also to instruments of a rectangular construction. In this case we get two sets of constants of normalization, namely $a = \frac{R}{l}$, $b = \frac{l}{r}$ and $a_1 = \frac{R_1}{r}$, and $b_1 = \frac{l_1}{r}$.

For characterizing the aperture conditions we get two aperture angles defined by:

$$\operatorname{tg} \frac{\varphi}{2} = \frac{a}{b} \text{ and } \operatorname{tg} \frac{\varphi_1}{2} = \frac{a_1}{b_1} \qquad (7)$$

A computation of the radiation falling on the rectangular receiver leads to rather complicated formulas, if great accuracy is attempted. (See Bossy and Pastiels, loc. cit.)

2. When we consider the case of the *pyrheliometer* and ask how the readings are related to the energy falling on the receiver and to the radiation which we intend to measure, account must be taken not only of the aperture conditions, the dimensions and sensitivity function of the receiver, but also of the fact that the radiation to be measured is not uniformly distributed over the space covered by the aperture. The circumsolar sky covered by the aperture has naturally a much smaller luminosity than the sun's disk, but on the other hand it extends for most pyrheliometers over a solid angle about 50 times larger than the solar disk itself. The luminosity of the circumsolar sky has been subjected to a number of studies, among which here may be mentioned those of Dorno and Thilenius, of F. Linke and Dubois, of F. Linke and E. Ulmitz, of D. Stranz, of F. Volz, E. V. Piaskovskaia-Fenkova and of L. Bossy and R. Pastiels. Linke and Ulmitz (1940) expressed the radiation from the circumsolar sky through the equation:

$$I = kE_0 e^{-c\alpha} \qquad (8)$$

where E_0 is the intensity at the center of the sun and α is the distance from the center in degrees. For the constant k Linke and Ulmitz gave a value which according to Bossy and Pastiels (1948) is about 50 times too high, due to an error in the calculation. Using the measurements of Linke and Ulmitz, Bossy and Pastiels derive for k a value of 0.000479 (instead of 0.0235) and thus arrive at the formula:

$$I = 0.00048 \cdot E_0 e^{-0.58\,\alpha} \qquad (9)$$

On the basis of the formula (8) and with introduction of the constants of formula (9), Bossy and Pastiels have computed the energy of the circumsolar radiation entering the pyrheliometers of the Smithsonian Institution (Silver Disk) and of Linke-Feussner. In this computation they have taken into consideration also the sensitivity distribution of the receiver. The introduction of the last-named factor gives results which generally deviate with up to about 10 per cent from the value for the circumsolar radiation obtained under the assumption that the receiving element is concentrated in a point (a point-receiver). Now, as we will discuss further below, the formulas (8) and (9) give only a very approximate idea of the distribution of the circumsolar radiation. Not only are, as already was emphasized by Linke and Ulmitz, the constants k and c variable to the extent that they in special cases are more than the double or less than the half of those presented, but also the form of the equation (8) itself represents an ideal case, which only seldom is satisfied. In many cases it ought to be supplemented by a factor, showing maxima and minima with the distance from the sun.

It seems therefore justified, for practical purposes, to simplify the computations and to base a general consideration of the effects of aperture conditions on the simple case of a point receiver situated within pyrheliometers of different apertures. For the conclusions aimed at *here* such a limitation seems fully satisfactory.

The total circumsolar sky radiation D falling on a point receiver, the aperture being $\varphi = 2\alpha$, is given by:

$$D = 2\pi \int_{n=1}^{n} Ir dr \cos \alpha \qquad (10)$$

where r is the distance from the center of the sun, with the radius of the sun as unit. For the small angles α here in question ($\alpha < 5°$), we may put $r = 3.72\,\alpha$. A further simplification, here justified, is arrived at by replacing cos α

by 1.00. Both these approximations introduce an error less than one per cent in D, and are justified with regard to the fact that the constants of equations (9), as we will see, must be subjected to rather great variations from case to case.

An integration of (10), after these approximations, gives:

$$D = 0.0030 \left[\frac{e^{-\gamma}}{\gamma} \left(1 + \frac{1}{\gamma} \right) - \frac{e^{-\gamma r}}{\gamma} \left(r + \frac{1}{\gamma} \right) \right] E_0 \quad (11)$$

As the radiation Q from the whole solar disk, on the basis of values available, may be approximated to $0.8\,\pi E_0$, we get for the ratio: $\varepsilon = \dfrac{D}{Q}$

$$\varepsilon = 0.0012 \left[\frac{e^{-\gamma}}{\gamma} \left(1 + \frac{1}{\gamma} \right) - \frac{e^{-\gamma r}}{\gamma} \left(r + \frac{1}{\gamma} \right) \right] \quad (12)$$

Here γ has, in the case corresponding to the observations of Linke and Ulmitz, a value:

$$\gamma = \frac{c}{3.72} = \frac{0.58}{3.72} = 0.156$$

The following table gives the value of ε for different values of φ, $\left(\varphi = 2\alpha = 2 \cdot \dfrac{r}{3.72} \right)$, and for three different values of γ:

Table 1

φ / γ	0	2°	4°	6°	8°
0.100	0	0.0050	0.020	0.035	0.051
0.156	0	0.0045	0.015	0.024	0.032
0.200	0	0.0040	0.013	0.019	0.023

Graphically the dependence of ε on the aperture angle (2α), in the case of a point receiver, is shown in fig. 2.

It is evident already from a superficial consideration of the process of scattering that the equation (9) can be valid only within a rather limited range. With the constants given, the equation (9) represents, as was emphasized by Linke and Ulmitz, mean values, from which the individual values deviate down to the half or up to the double. Linke and Ulmitz found for the ratio $\dfrac{D}{Q}$ values which were related to the product of air mass and turbidity (T) by an approximately linear relationship (for $\alpha < 5°$). This result has been confirmed by Volz, who also, as it is to be expected, finds an increase of ε with a decrease in wavelength. The turbidity factor T, which includes also the effect of water vapor absorption on the extinction, is not, however, a satisfactory measure of the scattering power of the atmosphere. W. Schüepp (1949) has treated statistically a rather large mass of observations from Davos taken (1) with the aid of a pyrheliograph with about 2° aperture and (2) by means of a Michelson actinometer with a rectangular opening of 6° × 15°. On the basis of comparisons between both sets of values, Schüepp derives an expression for the percentage difference, given as a function of the product of the turbidity coefficient and the air mass. Translating from the turbidity coefficient B of Schüepp to the coefficient β, employed in the present paper, we obtain:

$$\varepsilon = 0.315\, m\, \beta \quad (13)$$

where m is the air mass. Assuming $m = 1.2$ and $\beta = 0.08$, which may reasonably apply to the average conditions under which the measurements of Linke and Ulmitz were made, we obtain $\varepsilon = 3.0$ per cent, a value which comes very near to the 2.8 per cent derived from the diagram (fig. 2) for the difference corresponding to actinometers with apertures of 8° and 2° respectively.

Kalitin (1933 and 1937) found that the sky luminosity was not subjected to a continuous decrease with an increasing distance from the sun, but showed minima at about 2°.5 and 4°.5 distance; Linke and Ulmitz (1940) found

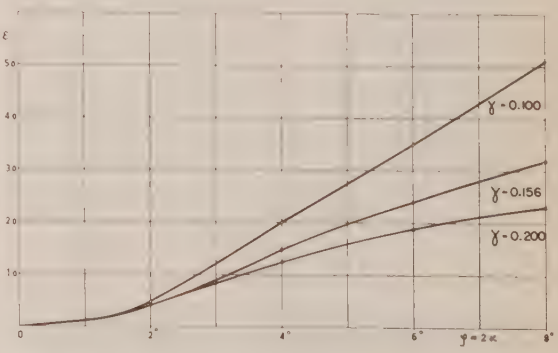


Fig. 2. Percentage increase, ε, of observed solar radiation, if measured by means of pyrheliometers (point receivers) with different apertures, φ. The coefficient γ is a measure of the decrease of circumsolar sky radiation with the distance from the sun.

Table 2. Apertures of pyrheliometers.

Pyrheliometer	aperture[1]
K. Ångström comp. pyrh.	5° × 13°
Silverdisk	6°
Michelson	5° × 13°
Moll-Gorzynski	8°
Linke-Feussner	11°
Eppley normal incid.	5·7°

[1] From IGY Instruction Manual, Vol. V, Part VI, Radiation Instruments and Measurements, Pergamon Press. 1958.

the decrease to turn into an increase at 6°.5 distance from the sun. This is in conformity also with the results of D. STRANZ (1940, 1942), who states that the scattered radiation from the sky, in the neighbourhood of the sun, is not, under many conditions, a simple decreasing function of the angular distance, but shows secondary maxima and minima, the location of which are dependent mainly on the nature and size of the scattering particles of the aerosole.

Recently G. A. NEWKIRK (1956) has made a photometric study of the solar aureole. No integral or absolute values are directly derived for the circumsolar radiation, but it may be concluded from his discussion, that, accepting the empirical formula (8), we must assume k as well as c to be dependent of the amount and character of the scattering particles of the atmosphere.

The conclusion must be that the values for the circumsolar sky radiation given in table 1 and fig. 2, can only represent average conditions, limited to a clear sky, a moderate or low turbidity, and a solar elevation corresponding to relative air masses less than about 2.5. At a solar elevation of, for instance 15°, the circumsolar sky radiation entering an actinometer of the type used in practical field work, may easily rise to about 10 per cent or even more.

3. Practical Considerations

It may be concluded from the previous discussion that, in the operation of the common pyrheliometers, K. Ångström, Silverdisk, Linke-Feussner, Eppley, Michelson, etc., we are never perfectly free from the effect of an additional sky radiation. Practically we may get rid of it by making the aperture less than about one degree, but such a construction has very evident practical drawbacks and we then must consider more closely what is to be gained by making the aperture very small.

a. *Solar constant determinations*

As the amount of circumsolar sky radiation is rather small and approaches zero in a linear way when the air mass decreases to zero, it seems evident that it can not be expected to influence to an appreciable degree the solar constant value, obtained through the so-called "long method". As this method is the basis also of the "short method" of Abbot it seems reasonable to assume that aperture effects do not introduce an appreciable error into solar constant determinations.

The error, which may sometimes occur from this cause, must in any rate be small compared with the error occasionally due to *variations in the atmospheric transmission* during the time covered by the set of radiation measurements. Such a change in the transmission must evidently affect the variation of Q as well as D and leads, when the values are extrapolated to air mass o, to an error dependent on the variation of both the named quantities. Under conditions of a very clear atmosphere, as on high mountain tops in dry climates, and in cases where the aperture is smaller than 8°, the named error probably seldom amounts to more than a small fraction of 1 per cent.

b. *Standard scales*

As regards the difference between the standard scales, the Ångström and the Smithsonian, a slight effect of differences in aperture conditions may reasonably be suspected. On the other hand no influence of the air mass or solar elevation on the difference of the scales has, so far, been shown to exist, which ought to have been the case if the difference of the scales was due mainly to this cause. P. COURVOISIER (1957) has presented a fundamental discussion on this question. It seems therefore probable, that for the average distribution of circumsolar sky radiation, the effective surface angle ratios of the pyrheliometers, which have represented the scales, are so closely equal as to eliminate differences occurring on this ac-

count[1]. It seems more probable that the variations of the difference between the two scales, which undoubtedly have been obtained on different occasions (about 3.5 ± 1.0 per cent), are due to other causes (Ångström 1958).

If we with the true pyrheliometric scale understand the values Q_0 given by an exact pyrheliometer, which measures the direct sun radiation only, without any additional sky radiation, we must realize, that there is no possibility whatsoever to derive a strict evaluation of this scale simply from measurements with pyrheliometers characterized by aperture angles like those generally used in practical work (Ångström, Silverdisk, Michelson, etc. pyrheliometers). The measured value is here always given by:

$$Q = Q_0 + D(m, \beta, \lambda, \varphi) \qquad (14)$$

where the circumsolar term D is a function of air mass m, turbidity β, wavelength λ and aperture φ, as the chief variables. We may guess at the value of D from measurements like those of Linke and Ulmitz, but as long as D is not directly measured at the instant of observing Q, and as the function D is not accurately known, all speculations as regards the differences between observed values and those of the true solar radiation itself are rather fruitless, at least if they aim at an accuracy greater than $\pm$ 1.0 per cent.

If we observe the radiation with, let us say, an Ångström pyrheliometer or a silver disk, the true value of the solar radiation *alone* may be between 1.0 and up to 5.0 per cent lower, dependent upon atmospheric and other conditions determining the variables in equation (14). Comparing two pyrheliometers characterized by different apertures, they may thus differ with one value today and another value tomorrow dependent upon the named variables. Strictly, therefore the true pyrheliometric scale value can never be given through applying a fixed constant percentage correction to the values obtained with pyrheliometers with aperture angles larger than a couple of degrees. Practical average values for the correction may however be derived. For greater accuracy, when in rare cases a need for such occurs, we are dependent on a clearer knowledge of the function D, or preferably on measurements with more specialized instruments.

[1] In the case of the K. Ångström compensation pyrheliometer the circumsolar sky radiation falling on the strip illuminated by the sun is partly compensated by the sky radiation reaching the screened strip. In view of the occurrence of secondary maxima, as mentioned above, it seems even possible that under extreme conditions an overcompensation may take place.

c. *Practical use*

In discussing the aperture effect and the question in regard to the desirability of its elimination, we must consider also the practical use, which the measurements are intended to serve. For many purposes, for instance when it comes to computing the heating effect on differently oriented surfaces, we wish to distinguish between the *directed radiation* from the sun, and the *diffuse radiation from the sky*, both groups differing fundamentally as regards their dependence on geometrical optical factors. For most practical evaluations it seems here reasonable to include a certain amount of circumsolar sky radiation in the solar radiation itself, as has been done in practice using pyrheliometers with apertures about 5° in size.

The circumsolar radiation naturally follows the sun and obeys the cosine law in its dependence on solar elevation and azimuth.

Acknowledgements

The present paper has been prepared under the Eppley Foundation for Research. For ample conveniences offered by the Eppley Laboratory the author is indebted to Mr Roy Anderson, president. He is also obliged to Mr A. J. Drummond, M. Sc., Chief Research Physicist of the Laboratory, and to Dr P. R. Gast for their contribution to many useful discussions on the matters here treated.

Appendix

In a report of the Sub-commission on Aperture conditions of the Radiation Commission of the International Association of Meteorology (IUGG) submitted to the Davos session of the Radiation Commission, in September 1956, the following parameters were recommended to specify the aperture conditions of a radiometer. (Compare fig. 1 and Linke's "Constants of normalization".)

α the "opening angle". $(\operatorname{tg}\alpha = \frac{a}{b})$

α_1 the "slope angle". $(\operatorname{tg}\alpha_1 = \frac{a-1}{b})$

α_2 the "limit angle" $(\operatorname{tg}\alpha_2 = \frac{a+1}{b})$

Evidently all these parameters can be derived from the constants of normalization, a and b.

The subcommission realized that it would be unpractical to recommend an aperture angle so small that the circumsolar radiation falling on the receiver would be practically negligible, but regarded it desirable, in order that measurements with different standard instruments should be comparable with one another, that a certain normalization of the aperture conditions were established. For this reason it was recommended that "pyrheliometers constructed in future should have an angle of slope not less than 1°, and not greater than 2°, and a ratio $\frac{l}{r}$ equal to or greater than 15. These conditions imply that the opening angle should not be greater than 4° ($\varphi < 8°$).

As a practical instance of the extent to which some pyrheliometers with circular apertures conform to these rules, we may refer to the following table:

Abbot Silver disk pyrheliometer:

R = 18.5 mm		α = 2°.9	
r = 13.5 »		α_1 = 0°.8	$b = 27$
l = 370 »		α_2 = 4°.9	

Eppley standard pyrheliometer:

R = 10.3 mm		α = 2°.9	
r = 2.8 »		α_1 = 2°.1	$b = 73$
l = 206 »		α_2 = 3°.6	

Pyrheliometer Linke-Feussner:

R = 6.3 mm		α = 5°.1	
r = 5.0 »		α_1 = 1°.0	$b = 14$
l = 70.3 »		α_2 = 9°.1	

It will be noticed that the geometry of the pyrheliometers are in reasonable agreement with the named requirements. It is especially important that the aperture angle ($2 \times 2.9°$) for both the first named pyrheliometers be well below the limit value 8°, and yet above about 4° ($\alpha > 2°$), which seems to represent a reasonable lower limit, when we are concerned with practical field observations within the meteorological station network.

REFERENCES

Ångström, A., 1958: On pyrheliometric measurements, *Tellus*, **10**, pp. 342—354.

Bossy, L., Pastiels, R., 1948: Etude des propriétés fondamentales des actinomètres. *Institut Royal Météorologique de Belgique. Mémoires*, XXIX.

Courvoisier, P., 1957: Contributions to Standard Pyrheliometry with an introduction by A. Ångström and W. Mörikofer, *Sveriges Meteor. & Hydrological Institute Communications, Series* **B**. No. 13.

Kalitin, N. N., 1933: De la radiation de la voute celeste autour du soleil pour les ondes longues et pour les ondes courtes. *Veröff. d. Sekretariats d. Intern. Meteor. Org.* **15**.

— 1937: Des erreurs pendant le mesurage de l'intensité de la radiation solaire importées par les nuages dans l'angle d'ouverture de l'actinomètre. *Secretariat de l'Organisation Meteor. Internationale*, **33**.

Lambert, J. H., 1760: *Photometria sive de mensura et gradibus luminis colorum et umbrae.*

Linke, F., 1931: Die Bedeutung des Öffnungsverhältnisses eines Aktinometers für Messungen der Sonnen- und Himmelsstrahlung. *Strahlentherapie*, **39**, p. 351.

Linke, F., Ulmitz, E., 1940: Messungen der zircumsolaren Himmelsstrahlung. *Meteor. Zscht*, **57**, p. 372.

Newkirk, G. A., Jr., 1956: Photometry of the Solar Aureole. *Journal of the Opt. Soc. of America*, **46**, 12.

Pastiels, R., 1959: Contribution à l'etude du problème des méthodes actinométriques. *Inst. Roy. Meteor. de Belgique, Publications Serie* **A**, 11.

Schuepp, W., 1949: *Die Bestimmung der Komponenten der atmosphärischen Trübung aus Aktinometermessungen.* Inaug. Diss. Wien, pp. 320—321.

Stranz, D., 1942: Die Himmelshelligkeit in unmittelbarer Nähe der Sonne. *Meteor. Zscht*, **59**, p. 120.
— 1940: Untersuchungen der Himmelshelligkeit in der Umgebung der Sonne *Veröff. Geophys. Inst. Univ. Leipzig*, **XII**, 4.

Volz, F., 1956: Report to the Radiation Symposium at Davos (mimeographed).

Rapid Geomagnetic Fluctuations

By STIG HARALDSON, Uppsala[1]

(Manuscript received June 4, 1960, revised version April 19, 1961.)

Abstract

The article presents an investigation of VLF-oscillations (0.2—200 c/s) in the earth's magnetic field. The measurements were made in the archipelago of Stockholm (59° N) during the period July 1957—June 1959. The oscillations were recorded by two different instruments: a high-speed Heiland oscillograph and a low-speed Esterline-Angus recorder. In the Esterline-Angus diagrams a morning minimum at about 8 o'clock local time was found. In the Heiland oscillograms three different features were found: 1. Frequent occurrence of sinusoidal fluctuations presenting a few very narrow bands; 2. Oscillations in the H- and Z-component are sometimes simultaneous, sometimes not; 3. The amplitudes in the Z-component are smaller than in the H-component. It has not been possible to find a correlation between the fluctuations investigated here and other geophysical phenomena. A possible mechanism for generation of the sinusoidal oscillations is discussed.

1. Introduction

Common geomagnetic variometers have an upper frequency limit about 0.2—0.5 c/s and more rapid fluctuations in the earth's magnetic field have not been recorded until recently, WILLIS (1948), AARONS (1951) etc. The investigation to be described here was started in the spring of 1957, its purpose being to measure the amplitudes in the frequency range 0.2—200 c/s.

2. The apparatus

Two coils were used in recording the fluctuations, one with 1,080 turns and a total area of 9,100 m^2 for the Z-component and another with 3,000 turns and an area of 7,100 m^2 for the H-component.

The resonance frequency of the coils appeared at about 170 and 25 c/s respectively and it was therefore not possible to calculate the field amplitudes directly from the voltages appearing across the terminals. However, it did prove possible to calibrate the apparatus directly in $m\gamma$/mm ($m\gamma$ = milligamma = 10^{-8} gauss) through introducing a known alternating current into a calibration winding surrounding each coil.

During a first period only one amplifier (two cascade connected, battery operated Tektronix Preamplifiers) was used. The H- and Z-components were recorded alternatively on a battery operated Heiland oscillograph. Because of the high paperspeed (7 cm/sec) recordings were made only during intervals of 1—2 minutes every hour.

[1] This investigation has been sponsored by the Swedish Research Institute of National Defence.

In order to eliminate the peak at about 25 c/s in the sensitivity of the H coil, damping was arranged for by applying a resistance of 2.2 kΩ across the coil and by using in the oscillograph a galvanometer with a resonance frequency of 150 c/s. The calibration curves for the apparatus are shown in fig. 1.

During the summer and autumn of 1958 the instrumentation was completed by addition of a modified Tektronix Preamplifier with an extra final stage. This made it possible to make records of both components simultaneously. At about the same time an Esterline-Angus recorder with low paperspeed, about 90 cm/hour, became available. It was then possible to make continuous recordings. As this instrument had only one channel, the Z-component was chosen.

Fig. 2 shows a block diagram of the equipment.

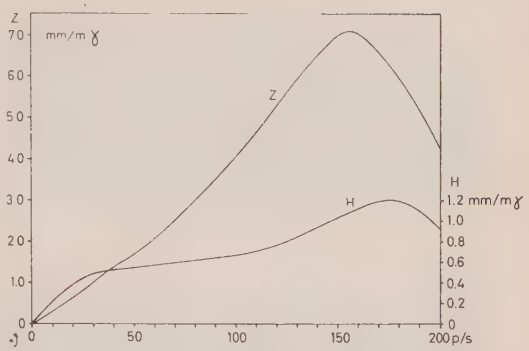


Fig. 1. Calibration curves. The sensitivity in mm mγ (millimeter per milligamma) as a function of the frequency.

3. Sites

The first measurements were made in the neighbourhood of Uppsala (60° N) during the spring 1957. All that was recorded was a 50 c/s hum from the AC power lines and a $16^2/_3$ c/s hum from the railroads. These disturbances did not decrease appreciably when the local transformer stations in the neighbourhood were disconnected even though the distance to the nearest main AC power transmission line was 2.5 km., the distance to the city of Uppsala 10 km and the distance to the nearest railroad 4 km. In order to eliminate sources of this type, the site had to be moved far away from inhabited regions, but for practical reasons could not be too far away from Uppsala and Stockholm. In July 1957 a test was made on a small island in the outer parts of the archipelago of Stockholm (59° N), about 15 km from the nearest small AC power line and 55 km from the nearest railroad. As no traces of 50 c/s or $16^2/_3$ c/s hum could be discovered, the island was chosen as site for the recordings.

The site had some great disadvantages, as the island was difficult of access, and the instruments were often badly treated in transit, each time recordings were to be made. Inorder to distribute the days of measurement as evenly as possible over the year, we decided to make recordings 2—4 days per month, but various misadventures during transportation to and from the site made it impossible to keep this programme and the information obtained is therefore rather imperfect from a statistical point of view. During the period 1. 7. 1958—30. 6. 1959, when the equipment was complete, there are only 14 days with recordings, that permit calculation of a fairly certain daily mean value.

The calibration also became a labourious procedure as the capacitance of the coils changed, due to condensation of moisture between the turns. The calibration had to be checked frequently. The noise level from the apparatus, which had to be very low, was controlled by connecting a resistance and a capacitance of the same order as the capacitance of the coils to the amplifier.

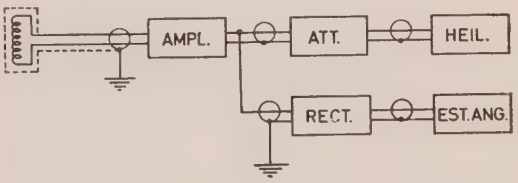


Fig. 2. Block diagram of the apparatus.

4. The Heiland oscillograms

A statistical analysis of the Heiland oscillograms has not yet been made and only some general and interesting features will be discussed. The most striking thing to be observed is that in addition to the noise there are not seldom variations similar to the giant pulsa-

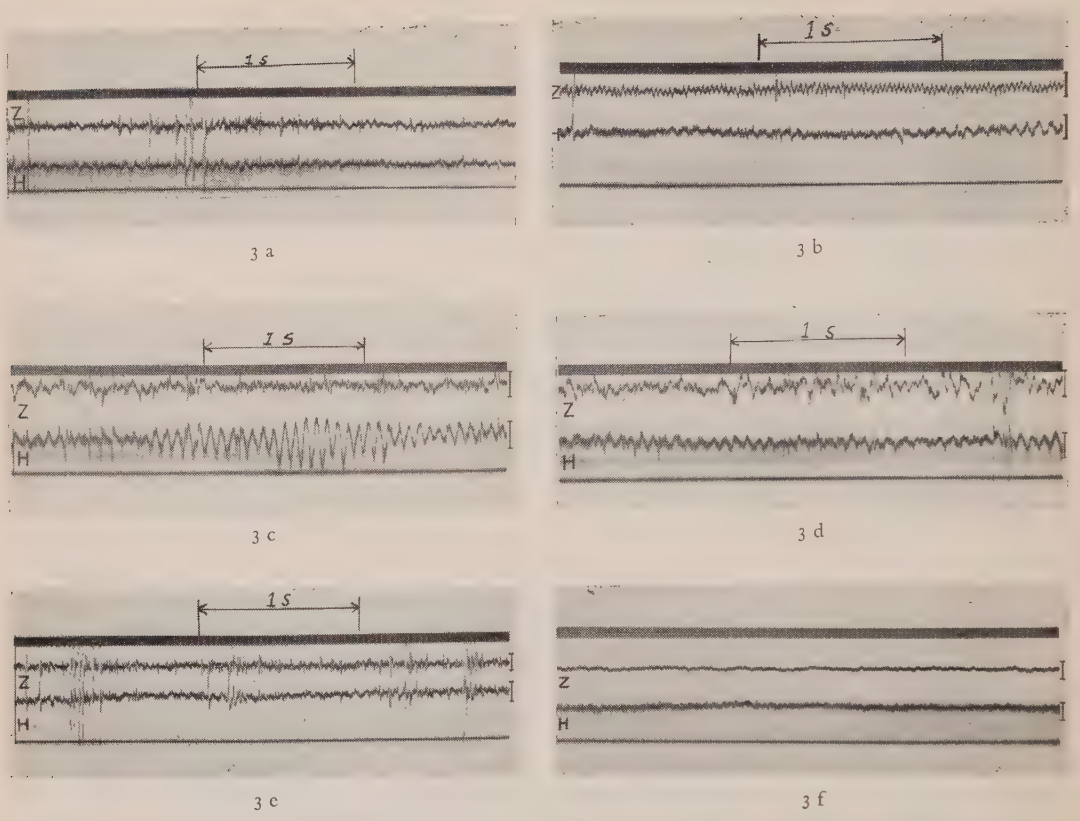


Fig. 3. Heiland oscillograms; a) 1.6. 1959 about 2119 o'clock GMT; b) 2. 6. 1959 about 0855 o'clock GMT; c) 2. 6. 1959 about 1228 GMT; d) 2.6. 1959 about 1229 GMT; e) 11. 6. 1959 about 1924 GMT; f) noise. The lines to the right of the curves give a measure of 10 mγ for the *Z*-component and 20 mγ for the *H*-component at 30 c/s.

tions, i.e. sinusoidal fluctuations with a very definite frequency, see fig. 3 a—e. Fig. 3 f shows the noise from the apparatus.

It is also remarkable that these fluctuations present only a few narrow bands, about 15 c/s, 20 c/s, 30 c/s, 40—50 c/s and 140—160 c/s. Sometimes, oscillations of a very low frequency, < 1 c/s can be found, but on account of the relatively low sensitivity for these frequencies they are rare.

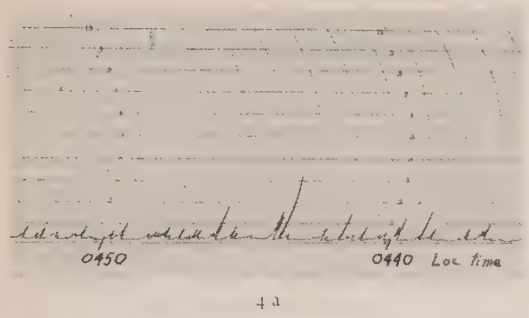
Another very astonishing feature is that the fluctuations in the H- and the Z-component are sometimes simultaneous, sometimes not, as is seen in the figures.

This was observed already before we could make simultaneous records of the H- and Z-component. In every record of H during one day 20 c/s fluctuations could e.g. be found, but no such oscillations appeared in Z. As the total recording time was rather short, the possibility of pure coincidence could not be left out. The simultaneous recordings made later on verified, however, this observation.

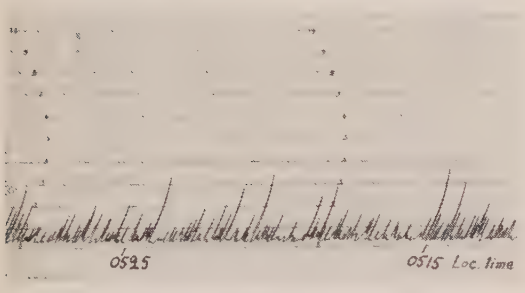
The third general feature is that the amplitudes in the Z-component on the whole are much smaller than in the H-component.

Examples of amplitudes:

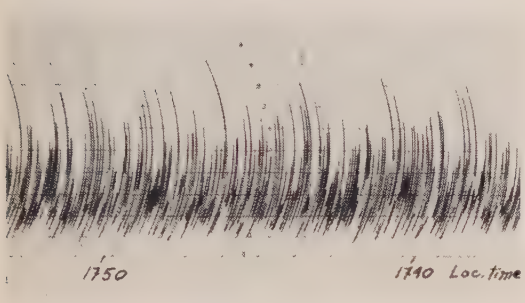
about 150 c/s:	H	10—20 mγ	up to 100—120 mγ
	Z	1—2 mγ	up to 10 mγ
30 c/s:	H	10—20 mγ	up to 150 mγ
	Z	1—3 mγ	up to 10 mγ
20 c/s:	H	15—20 mγ	up to 150 mγ
	Z	5 mγ	up to 30 mγ
< 1 c/s:	H		up to 30 γ
	Z		up to 2 γ



4 a



4 b



4 c

Fig. 4. Esterline—Angus diagrams: a) 9. 1. 1959, a quiet day; b) 21. 10. 1958, a moderately disturbed day; c) 11. 6. 1959, a very disturbed day.

5. Esterline-Angus diagram

Figs. 4 a—c show examples of Esterline-Angus diagrams for quiet days, moderately disturbed days and very disturbed days. As is seen in 4a and 4b, it is not so much the basic level nor the amplitude of the individual "spikes" that are affected, but the "spikes" are more frequent on a disturbed day. The diagrams were divided in half-hour intervals for purpose of analysis. For each half-hour the number of spikes ranging up to between 2.5—5 scale-divisions, 5—10 divisions, 10—15 divisions etc. was counted and the numbers multiplied by the lower limit-value for the interval. The disturbance figure for each half-hour was then defined as the sum of the products.

In order to find an eventual daily variation, the disturbance figure was then expressed in per cent of its daily mean value and the relative disturbance figures for the same half hour of all days were added. This calculation has been confined to the half day 0000—1200 local time, since the sequence of observation was more complete during this interval, gaps during 1200—2400 being due mostly to calibration checks.

As seen in fig. 5 there is a marked minimum about 0800 local time. This confirms what can be directly observed on the diagrams. It is also in agreement with almost all earlier investigation of VLF-noise in middle latitudes. The results reported by EGELAND (1959) and EGELAND et al (1960) indicate that conditions in the auroral zone (68° N, 70° N and 75° N) may be different.

The geomagnetic K-index for the near-by Lovö Geomagnetic Observatory and the disturbance figures reported here show the same average daily variation but it has not been possible to find a correlation between actual values. The same is true for the K_p-index. On this point earlier investigations have given different results. A correlation has been noted by AARONS (1951), EGELAND (1959), EGELAND et al (1960) and MAPLE (1960), but not by Aaron's in two papers from 1953 and 1954. It is here interesting to note that in almost all cases where a correlation has been reported, the measurements have been made in the auroral zone, the exception being AARONS (1951) and (1956). Thus the validity of this correlation and the extent of its latitude dependency are still uncertain.

It should also be noticed that the noise figures, as calculated above, represent a kind of mean value over the interval 0.2—200 c/s, with more weight attached to the higher frequencies and that at least some of the investigations mentioned indicate that the correlation may exist for frequencies below, say, 50 c/s, but not for higher frequencies. In further investigations one should therefore use an apparatus with several channels in order that the amplitudes could be examined over smaller frequency intervals.

We have also tried to correlate the disturbance figures with variations in the cosmic

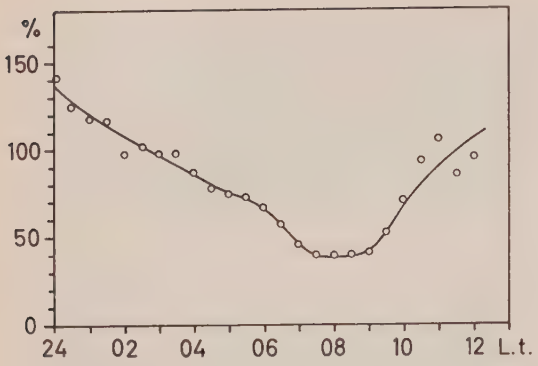


Fig. 5. Relative disturbance figures in per cent of the daily mean v. local time.

ray intensity. At first it seemed as if there was a connection of some kind between these two phenomena, perhaps with a phase shift of one day. Further investigation, however, showed that this connection was very obscure and it is impossible to deduce a correlation between our disturbance figures and the variations in the cosmic ray intensity, recorded at Uppsala.

Similarly, all efforts to find a correlation with ionospheric characteristics, such as the critical frequency for various ionospheric layers a.s.f., have failed.

Now these failures, of course, do not necessarily mean that there is no connection between our disturbance figures and the other geophysical phenomena mentioned. It is more reasonable to think that there is a very complex connection between a great many factors, which makes it very hard to put the things together. It must also be remembered that from a statistical point of view our material is rather small.

6. Origin of the fluctuations

Many of the fluctuations are very similar in form to the giant pulsations, although of much higher frequency. LEHNERT (1956) and (1959) has tried to explain the giant pulsations with magnetohydrodynamic waves in the ionosphere. His calculations show that for the frequencies investigated here, the damping will certainly be too great, so it seems impossible to explain these fluctuations in that way.

Aarons has suggested that ions spiralling in the earth's magnetic field could give rise to fluctuations with the gyrofrequency. It can be worth while considering this possibility. The process must take place at an altitude of 200 – 300 km in order that the mean free path be large enough for the ions to complete several rounds between the collisions.

According to the law of Biot-Savart, a charged particle with a velocity v in the near zone ($r \ll \lambda$) gives a magnetic field with a flux density

$$B = \frac{\mu_0}{4\pi} \cdot \frac{e\, v \times r}{r^3}$$

If we, as Aarons does, let v be the thermal velocity at 300 km altitude, we have $v \sim 6{,}000$ m/s. This can give an amplitude for $B \sim 10^{-33}$ Vs/m^2. To obtain a total amplitude of 10 mγ $= 10^{-11}$ Vs/m^2, we thus need totally 10^{22} ions revolving in a coherent manner. This does not seem very likely. On the other hand, if there is a much greater number spiralling at random, it is not quite impossible that there could occasionally exist a net vector sum of a detectable size. It is also possible, as Aarons has pointed out, that the energy of the ions is suddenly increased by e.g. meteor impacts.

The gyrofrequency is given by the formula

$$f = \frac{B \cdot e}{2\pi\, \mathrm{m}}$$

where B is the total geomagnetic flux density, in Stockholm 0.49 gs, say 0.45 gs and 0.44 gs at an altitude of 200 resp. 300 km.

Table 1 below gives the gyrofrequencies for some of the ions in the ionosphere.

These figures are very suggestive because the frequencies of the sinusoidal fluctuations seen in the Heiland oscillograms almost always

Table 1

	f in c/s at height	
ion	200 km	300 km
H^+	690	680 km
He^+	170	170
N^+	49	48
N_2^+	25	24
O^+	43	42
O_2^+	22	21
O_3^+	14	14
Na^+	30	30

lie around 20 c/s, 30 c/s, 40—50 c/s. The theory is also consistent with the fact that the fluctuations are more frequent in the H- than in the Z-component. Of course, ions rotating over our heads are nearer and therefore easier to detect than ions on a great angular distance from our zenith-direction. But the field from the ions overhead has no or only a small component in the Z-direction, which will be parallel to r.

It would be very interesting to record simultaneously fluctuations in two rectangular horizontal directions to see if the fluctuations have a phase shift of 90° in these two components.

This theory, however, is still only guesswork and we are far from a true knowledge about these fluctuations and their origin.

Acknowledgement

The author is deeply indepted to Laborator Sigfrid Wennerberg, Research Institute of National Defence, for support and encouragement during the whole progress of the work. Thanks are also due to Professor Per Ohlin and Docent Einar Lyttkens, Uppsala and Laborator Nicolai Herlofsson, Stockholm, for valuable advice and criticism.

REFERENCES

AARONS, J., 1951: Low-Frequency Noise Disturbance, *Proc. Am. Acad. Arts and Sc.*, **79**, p. 266.

AARONS, J. and HENISSART, M., 1953: Low Frequency Noise in the Range 0.5—20 c/s, *Nature*, **172**, p. 682.

AARONS, J., 1954: *Doctoral Thesis*, University of Paris.

— 1956: Low-frequency Electromagnetic Radiation 10—900 c/s, *J. Geophys. Res.* **61**, p. 647.

AARONS, J., GUSTAFSSON, G. and EGELAND, A., 1960: Correlation of Audio-Frequency Electro-magnetic Radiation with Auroral Zone Micropulsations, *Nature*, **185**, p. 148.

— 1960: Audio-Frequency Electromagnetic Radiation in the Auroral Zone, *J. Geophys. Res.* **65**, p. 2749.

EGELAND, A., 1959: Report No. 6. Inst. of Theoret. Astrophysics, Blindern, Oslo.

LEHNERT, B., 1956: Magneto-hydrodynamic Waves in the Ionosphere and their Application to Giant Pulsations, *Tellus*, **8**, p. 241.

— 1959: Plasma Physics on Cosmical and Laboratory Scale, *Nuovo Cimento*, Suppl. XIII, Ser. X.

MAPLE, E., 1960: *Sub-audio-frequency (1 to 50 c/s) Geomagnetic Fluctuations*, presented at the Geophysical Conference, Helsinki, July—August 1960.

MENZEL and SALISBURY, 1948: Audio Frequency Radio Waves from the Sun, *Nature*, **161**, p. 91.

WILLIS, H., 1948: Audio Frequency Magnetic Fluctuations, *Nature*, **161**, p. 887.

SHORTER CONTRIBUTION

A Note on the Diurnal Wind Variation

By G. J. HALTINER, U,S. Naval Postgraduate School, Monterey, Calif.

(Manuscript received May 4, 1961)

In recent correspondence HARRIS (1960) suggested that the discrepancy between observed winds and those computed from a theoretical model for the diurnal wind variation (HALTINER 1959) may be due, at least in part, to the diurnal pressure variation which had been omitted in the theoretical model. In a reply (1960) the author provided an analytical solution for the diurnal wind variation resulting from a diurnal pressure variation with frictionless flow. The solution indicated that the diurnal wind variations due to corresponding pressure variations were similar to inertial oscillations, though a different period. With respect to the latter, it is found that where the friction force is large, inertial oscillations initiated by geostrophic deviations, other than those caused by friction, usually are quickly damped. Hence it might be expected that wind oscillations due to pressure variations would be suppressed in a similar manner. Nevertheless actual solutions were desirable to verify this inference and the purpose of this note is to report such results.

In addition it was desired to test a form for the coefficient of eddy viscosity suggested by some results of BLACKADAR (1959) who used the geostrophic departure method to determine the vertical and diurnal distribution of the eddy viscosity from a set of averaged wind observations for a 24 day-period. A harmonic analysis of the data (MCBRIDE 1960) showed that about six harmonics were required to fit the data accurately.

However, the general form of the distribution of eddy viscosity can be represented by the mean value as a function of height plus the first harmonic in which the amplitude and phase are also functions of height as follows:

$$K(z,t) = K_0\{10|z-5|^2 + 200 + [8(z-5)^2 + 175]\cos\left[\frac{2\pi}{24}t + \frac{(21-1.2z)\pi}{180}\right]\} \qquad (1)$$

Here the height z is in hundreds of meters, the time t in hrs, and K_0 is a constant governing the over all magnitude of the eddy viscosity. For a value of K_0 of 400 cm^2/sec, the following table indicates the general range of the eddy viscosity as represented by expression (1).

Height	2400 meters	500 meters	Surface
Maximum K(cm^2/sec)	2.75×10^6	1.5×10^5	3.3×10^5
Minimum K	3×10^5	10^4	3×10^4

In addition, values of K_0 of 800, 200, and 100 cm^2/sec were also used in the numerical integrations in order to determine if such differences in the mean magnitude of eddy viscosity made a significant change in the wind pattern.

The equations of motion which were integrated may be expressed as

Table 1

CASE	u_{min}(m/s)	t_{min}(C)	Alt(m)	u_{max}(m/s)	t_{max}(C)	Alt(m)	v_{max}(m/s)	t_{max}(C)	Alt(m)
A	— 2.2	0100C	200	2.5	0300C	800	9	0300C	425
B		similar to Case A							
C	— 2	0200C	240	1	0400C	50	8.2	0300C	600
D	— 1.35	1700C	240	3.5	2100C	800	9.2	2000C	600
E	— 1.6	0200C	240	2.8	0500C	800	10.7	0400C	600
F	— 1.4	1700C	240	5.0	2300C	800	10.5	2200C	600
Observed	— 3.5	0200C	300	5.5	0600C	700	15	0100C	500

* Only in Cases A and B was the integration performed for a period exceeding 24 hrs; in the other cases integration was stopped efter the maximum value of v had been reached.

$$\frac{\partial u}{\partial t} = f(v - \bar{v}_g) - \frac{1}{\varrho}\frac{\partial p'}{\partial x} + \frac{\partial}{\partial z}\left(K\frac{\partial u}{\partial z}\right)$$
$$\frac{\partial v}{\partial t} = -f(u - \bar{u}_g) + \frac{\partial}{\partial z}\left(K\frac{\partial v}{\partial z}\right) \qquad (2)$$

Here p' represents the periodic part of the pressure; and $f\bar{u}_g$ & $f\bar{v}_g$, which are functions of height, give the mean pressure force with respect to time. The numerical integration method was the same as used earlier. The upper boundary of the layer was taken to be 2400 meters and the vertical height increment was 60 meters. A time increment of about two minutes was required to maintain computational stability.

The equations were integrated with (Case A) and without (Case B) the periodic pressure term, as well as with the other values of K_0, namely, 800 cm²/sec (Case C), 100 cm²/sec (Case D). Two additional cases were tried with K_0 = 800 cm²/sec (Case E) and K_0 = 200 cm²/sec (Case F), in which some indirectly determined values of the geostrophic wind were used instead of the measured values. The former were determined in such a way so as to minimize differences between the computed values of the two coefficients of eddy viscosity individually calculated from the x and y-equations of motion. Table 1 indicates the main results by giving the maximum and minimum values of computed and observed wind components, together with their times of occurrence.

It may be concluded from these experiments that the term representing the diurnal pressure variations made no significant difference in the diurnal wind variation in the friction layer. Secondly, the function (1) for the coefficient of eddy viscosity gives a fair representation of the diurnal wind variation for the period investigated. The largest source of error appears to be in the maximum magnitudes of the velocity components which are only from about $^1/_2$ to $^2/_3$ the observed values. Doubling the basic magnitude of the eddy viscosity did not appear to make a pronounced difference in the results, but reducing it by a factor of two or more gave rather marked differences, particularly in the time of occurrence of maxima and minima.

REFERENCES

Blackadar, A. J., 1959: Final report on study of forecasting low-level wind gradients. AFCRC-TR-59-220, Pennsylvania State University.

Haltiner, G. J., 1959: The diurnal variation of the wind, *Tellus*, **11**, pp. 452—458.

Haltiner, G. J., 1960: Correspondence to the Editor, *Tellus*, **12**.

Harris, M. F., 1960: Correspondence to the Editor *Tellus*, **12**.

McBride, J. R., 1960: Determination of the coefficient of eddy viscosity as a function of height and time, etc., U.S. Naval Postgraduate School Thesis.